

PRECISION MEASUREMENT OF THE HIGGS BOSON MASS AND SEARCH FOR
DILEPTON MASS RESONANCES IN $H \rightarrow 4\ell$ DECAYS USING THE CMS DETECTOR AT
THE LHC

By

JAKE ROSENZWEIG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2022

© 2022 Jake Rosenzweig

I dedicate this to Jacob Myhre.

ACKNOWLEDGEMENTS

Without so many two- and four-legged blessings along the way, I could never have made it to this point in my academic career. Thus, I give my infinite thanks to the following folks.

To my high energy physics mentors, Professors Andrey Korytov and Guenkah Mitselmakher, for granting me this one-of-a-kind opportunity to do real *science* at CERN.

To my wife, Suzanne Rosenzweig, for showing me that dreams do come true. To my mother and father, Vicki and John, who always reassured me that I could achieve anything I put my mind to. Sleep peacefully, Mom. To my siblings, Alex, Ryan, Devin, Jace, and Claudia who frequently and gently reminded me that there was life outside of grad school.

Aunt Rach, Uncle Yuri,

To my mentor, Sheldon Friedman, and his wife, Rita Friedman (Rosenzweig), who chose to invest in my success at a young age. I have only made it this far thanks to your undying encouragement, love, and optimism. To Sheldon's best friend, Dr. Bernard Khoury, whose reputation has helped pave my road.

To Dr. Filippo Errico for his Dr. Lucien Lo To Dr. Noah Steinberg To Darin Acosta, for spending hours of discussion on every t To the gentle gents who introduced me to the world of CMS, Brendan Regnery and Bhargav Joshi.

To my comrades for showing me what it takes to survive the core courses, Dr. Atool Divakarla, Dr. Brien O'Brendan, Dr. Donyell Guerrero, and Dr. Vladinar Martinez.

To the many students who tagged along in our "CMS Office Hours": Sean Kent, Jeremiah Anglin, Cris Caballeros, Ari Gonzalez, Evan Koenig, Nik Menendez, Neha Rawal, John Rötter. And to the many students who let us practice our spiels:

To my mentee, Matthew Dittrich, for

To my Polish roommates in Saint-Genis-Pouilly for showing me what home away from home feels like.

To the many moms who generously gave unconditional support during the darkest times and unconditional love during the brightest times: Silet Wiley, Margaret Sherrill, Dawn Hood, Cyndi Reilly-Rogers.

To my childhood best friends: Jish, Willis, Shane, Zac, Duck, and Marcus for their constant

clever competition which has shaped me into the determined man I am today.

Big Tree:

And finally to Existence itself for this unpredictable, unbelievable blip of an experience called life.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGEMENTS	4
LIST OF TABLES.....	8
LIST OF FIGURES.....	11
CHAPTER	
1 INTRODUCTION	13
2 THEORY.....	14
2.1 The Standard Model	14
2.2 The Particles: The Players on the Fields.....	15
2.2.1 Bosons: Use the Force	17
2.2.2 Fermions: Each One Matters	19
2.3 Electroweak Symmetry Breaking and the Higgs Boson.....	20
2.4 The Standard Model Doesn't Explain Everything	21
2.4.1 Electroweak Interaction	23
2.4.2 Strong Interaction.....	23
2.4.3 Higgs Mechanism	23
2.5 Higgs Boson Production at the LHC	23
3 THE LARGE HADRON COLLIDER.....	24
4 THE CMS DETECTOR.....	25
4.1 The Silicon Tracker	28
4.1.1 The Pixel Detector.....	28
4.1.2 The Strip Detector	30
4.2 The Calorimeters	30
4.2.1 Electromagnetic Calorimeter	30
4.2.2 Hadron Calorimeter.....	33
4.3 The Solenoid and the Steel Return Yoke.....	35
4.4 The Muon System	37
4.4.1 Cathode Strip Chambers	38
4.4.2 Drift-Tube Chambers	43
4.4.3 Resistive Plate Chambers.....	46
4.4.4 Gas Electron Multipliers	47
5 THE COMPACT MUON SOLENOID EXPERIMENT	48
5.1 Trigger System	48
5.1.1 The Level-1 Trigger.....	48
5.1.2 The High-Level Trigger.....	49
5.2 Event Reconstruction	49
5.2.1 Track and Vertex Reconstruction	49
5.2.2 Electron and Photon Reconstruction	49

5.2.3	Jet Reconstruction	49
5.2.4	Muon Reconstruction	49
5.2.5	MET Reconstruction.....	49
5.2.6	Tau Reconstruction	49
5.3	Event Simulation	49
5.3.1	Hard Interactions	49
5.3.2	Parton Showering and Hadronization	49
5.3.3	Detector Simulation	49
6	HIGGS BOSON MASS MEASUREMENT IN THE $H \rightarrow ZZ^* \rightarrow 4\ell$ CHANNEL	50
6.1	Introduction	50
6.2	Analysis Overview.....	51
6.3	Data Sets, Simulated Samples, and Triggers	54
6.3.1	Data Sets.....	54
6.3.2	Simulated Samples	54
6.3.3	Triggers.....	54
6.4	Event Reconstruction and Selection.....	54
6.4.1	Event Reconstruction	54
6.4.2	Event Selection.....	54
6.5	Background estimation.....	54
6.5.1	Irreducible background	54
6.5.2	Reducible background	55
6.6	Signal Modeling	67
6.6.1	Z_1 Mass Constraint	67
6.6.2	Per-Event Relative Mass Error Categorization	67
6.6.3	VXBS Constraint.....	67
6.6.4	Matrix Element-Based Kinematic Discriminant	67
6.7	Systematic Uncertainties	69
6.8	Results	69
6.9	Summary.....	69
7	SEARCH FOR LOW-MASS DILEPTON RESONANCES IN THE $H \rightarrow ZX/XX \rightarrow 4\ell$ CHANNEL.....	70
7.1	Introduction	70
8	SUMMARY	71
	REFERENCES	71

LIST OF TABLES

<u>Tables</u>	<u>page</u>
6-1	67

LIST OF FIGURES

<u>Figures</u>	<u>page</u>
2-1 The fundamental particles of the Standard Model.	16
2-2 A Feynman diagram showing electron-electron scattering, also known as Møller scattering.	17
2-3 Two protons can be smashed together at very high energies to have their partons interact and convert the high energies into new kinds of matter.	22
4-1 Life-size poster of the CMS detector, taken during CERN Open Days 2019 in the SX5 warehouse where parts of CMS were assembled.	25
4-2 Points 1 through 8 along the LHC. Collisions occur at Points 1 (ATLAS), 2 (ALICE), 5 (CMS), and 8 (LHCb), whereas the remaining points are used for LHC beam maintenance and testing.	26
4-3 Cut out of the CMS detector showing its various subdetector components.	27
4-4 A transverse view of CMS showing the “filtration process” as different particles pass through different subdetectors. A positron (solid red line) curves due to the presence of the magnetic field and gets stopped in the ECAL, creating an EM shower. A photon (blue dashed line) does not get detected at all by the Silicon Tracker, since it has no electric charge. It continues through to the ECAL and makes a shower here, like the positron. Charged hadrons (solid green line) will show curved tracks from the Silicon Tracker, may leave some trace in the ECAL, but primarily get stopped by the HCAL creating hadronic showers. Neutral hadrons (dashed green line) do not interact with the tracker, and only undergo EM showers a little in the ECAL, but show most energy deposits in the HCAL. Muons (solid blue line) are detected by the Silicon Tracker and then mostly pass through the other subdetectors without interacting until they finally reach the Muon System. Using the Lorentz force law and knowing which direction the magnetic field is pointing, one can deduce the sign of the charge of the particle. Based on the radius of curvature from the trajectory, one can then calculate the momentum and energy of the particle.	29
4-5 (Left) A simulation of the silicon tracker, showing the 3 cylindrical layers of the pixel detector (pink), 4 layers of the TIB (green), and the 6 layers of the TOB (blue) of the strip detector. The endcap components are also shown. (Right) A picture of the real silicon tracker at the center of CMS.	30
4-6 A transverse view of the silicon pixel and strip detectors, explicitly labelling the different layers involved.	31
4-7 (Left) Cross sectional view of the electromagnetic calorimeter of CMS. (Right) One of the Dees which comprise the EE. Each square of 5×5 crystals constitutes a “supercrystal”. Figure taken from Ref. [1].	32
4-8 (Left) ECAL crystals made from PbWO ₄ are grown in a lab. (Right) Although made mostly of metal, ECAL crystals are transparent and have a photomultiplier detector attached at the end.....	32

4-9 A cross-sectional quadrant of CMS showing the locations of the HCAL components: the barrel (HB), outer (HO), endcap (HE), and forward (HF) detectors.	35
4-10 A longitudinal cross section of CMS showing the values of the magnetic field over the volume of CMS and various field lines. The magnetic field reaches its maximum of 3.8 T in the center of the detector.	36
4-11 (Left) A simulated cut out view of the ME+ endcap, with the coordinate system of CMS in the top-left corner. (Right) The actual ME-2 disk of CMS is shown, revealing its ME-2/1 and ME-2/2 rings of CSCs.	38
4-12 Longitudinal cross section of CMS, showing the different pseudorapidity values (η) and also the different subdetector regions.	39
4-13 Exploded view of the cross section of a CSC showing how the 7 panels come together to form the 6 gas gaps between the panels.	40
4-14 (Left) A CSC with its top layer exposed. You can see very thin gold-plated tungsten wires which actually span the entire width of the CSC. Thicker vertical strips run along the length. (Right) More detail showing the radial strips and the horizontal wires. Also shown is the 5 segments of a CSC.	41
4-15 A muon passes through one of the gaseous layers of the CSC, ionizing the gas mixture and inducing a charge on the anode wires and cathode strips.	42
4-16 (Left) Comparators are used to compare neighboring strip cluster charge to determine on which half-strip the peak charge resided. (Right) A muon passes through all six layers of a CSC inducing charge on various half-strips.	42
4-17 A drift-tube chamber with 3 superlayers (SLs) is sandwiching a honeycomb plate. Two SLs are indicated by the blue rectangles, whose cells (orange rectangle) have anode wires (red dots) oriented parallel to the beam pipe. The third SL is indicated by the green rectangle and is staggered orthogonally to the other two SLs.	44
4-18 A cross section of a DT cell showing the drift lines (light blue), isochrone lines (dark blue), dimensions of the cell, and locations of the anode wire and cathode strips.	45
4-19 A cross section of CMS showing the locations and numbering scheme of the drift tubes in the barrel.	45
6-1 The branching ratios of various Higgs boson decays as a function of the Higgs boson mass over a wide range of values (Left) and a narrow range (Right).	52
6-2 Contributions of per-event weights (boxed values) of various n -prompt-lepton processes ($X_{n\text{pr}}$, in parentheses) to the total numbers of events in the observed control regions ($N_{2\text{P}2\text{F}}, N_{3\text{P}1\text{F}}$) and signal region ($N_{4\text{P}}$). The labels ‘IB’ and ‘RB’ indicate those contributions which comprise the irreducible and reducible backgrounds, respectively.	59

6-3 Events from 2016 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, Z/γ^* , ZZ). The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).	61
6-4 Events from 2017 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, Z/γ^* , ZZ). The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).	62
6-5 Events from 2018 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, Z/γ^* , ZZ). The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).	63
6-6 Events from 2016 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, Z/γ^* , ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 6-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).	64
6-7 Events from 2017 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, Z/γ^* , ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 6-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).	65
6-8 Events from 2018 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, Z/γ^* , ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 6-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).	66
6-9 Electron (top row) and muon (bottom row) fake rates evaluated using the OS Method vs. the p_T of the probe lepton for 2016, 2017, and 2018 UL data (left, middle, right columns, respectively).	68

Abstract goes here.

Abstracts for doctoral dissertations must not be more than 350 words. Should be a concise summary of your dissertation's or thesis' study and purpose, fully understandable without reference to the dissertation or thesis itself. Your abstract ought not to contain any parenthetical or bracketed references.

CHAPTER 1 INTRODUCTION

The universe, while overwhelmingly vast, is comprised of a curiously small number of elementary particles. These particles and their strong, weak, and electromagnetic interactions with each other are accurately described by the Standard Model (SM). A major shortcoming of the SM was its inability to predict the masses of these particles.

This dissertation presents a precision measurement of the Higgs boson mass and using LHC proton-proton collision data from Run 2 data set from

The SM was not able to predict the masses of these particles until 1964 when the Brout-Englert-Higgs mechanism suggested that It wasn't until 1964 that the Brout-Englert-Higgs mechanism gave a self-consistent way to : by breaking the electroweak gauge symmetry of the vacuum would give rise to non-zero masses of the weak gauge bosons. This would yield a secondary effect too: there should exist a fundamental scalar boson which is the quantum of the so-called "Higgs field". On July 4th, 2012, this Higgs boson was discovered.

At first glance, the universe appears to be an overwhelmingly vast and complicated place. However upon closer inspection, it is comprised of only a few different kinds of fundamental particles. Particle physics has given rise to the Standard Model (SM) which mathematically describes these constituents and their interactions with each other.

The Standard Model (SM) is an impressively accurate mathematical theory which describes the fundamental particles of the universe and the rules for their possible interactions. Problematically though, the SM predicts that all particles are massless.

Get to the Higgs boson.

Why is it important? Knowing the mass of the Higgs boson

CHAPTER 2 THEORY

The Standard Model (SM) is a collection of the most accurate and self-consistent theories which mathematically describe the particles within the universe and their interactions with each other. For the past century, some of the most brilliant minds in physics have spent their entire careers to develop equations, mathematical tricks, and completely novel ideas to help build a solid foundation for the SM. To demonstrate the unparalleled accuracy with which the SM predicts , one can compare the experimentally measured anomalous magnetic moment of the electron

$$a_e^{\text{exp}} = 0.001\ 159\ 652\ 180\ 73(28)$$

to the value predicted by the SM

$$a_e^{\text{pred}} = 0.001\ 159\ 652\ 181\ 643(764).$$

An impressive agreement to better than one in a trillion!

In the following sections, the

First, the major players that make up the SM - the fundamental particles - are introduced. Next, an explanation is given on how the infamous Higgs boson completed the SM. Finally, problems of the SM are described and how its claims can be tested.

2.1 The Standard Model

The SM is a non-Abelian set of gauge theories. Non-Abelian gauge symmetry gives rise to charged gauge bosons.

$$SU(3)_C \times SU(2)_L \times U(1)_Y$$

Charge and hypercharge both obey U(1) symmetry.

Weak hypercharge, weak isospin (L reminds us that this interaction only effects left-handed fermions), color charge Q : *What is the difference between $U(1)_q$ and $U(1)_Y$? Q : Does L stand for left-handed?*

All observers must Lagrangian The connection between conservation laws and symmetries is best analyzed via Lagrangian mechanics.

Each Lagrangian has its own set of “Feynman rules”.

2.2 The Particles: The Players on the Fields

Contrary to intuition, fundamental particles are not hard, billiard-ball-like objects as is often perceived. Instead the SM predicts that every particle is actually an *excitation* of its corresponding field. So for example the electron is an excitation of the *electron field*, $\psi_e(x)$, a bispinor that describes a spin-1/2 field which follows the Dirac equation:

$$(i\hbar c \gamma^\mu \partial_\mu - mc^2) \psi(x) = 0.$$

In fact, *every* electron is an excitation of this same electron field. An electron is identical to every other electron in every way (same mass, same charge, etc.). Quantum Mechanics (QM) does a fine job of describing how slow-moving particles behave. However, as soon as particles begin moving at 30.5% of the speed of light (91.4 91.4E6 m/s), then there is a 5% difference between the particle’s rest mass energy ($E = mc^2$) and its total relativistic energy ($E = \gamma mc^2$). In other words, fast-moving particles must account for effects due to Special Relativity (SR). The merger of QM and SR gives rise to Quantum Field Theory (QFT) - the backbone of the SM.

Figure 2-1 shows all the fundamental particles that have been discovered up to the present day. The phrase “fundamental particle” just means that the particle is not composed of anything smaller than itself. These particles are not just diabolical creations from theorists. No, these particles are precisely defined, mathematical objects whose existence has been predicted by the SM and experimentally verified time and time again in the laboratory.

Each particle has a unique set of properties (like mass, electric charge, spin, etc..) that distinguish it from all the other particles. It is one of the primary goals of particle physics to determine these properties, because ultimately their properties determine their *interactions* with one another.

Without further ado, let’s meet the particles. There are two major types: *bosons* and *fermions*. We are going to take a non-traditional route and introduce the bosons first, then the fermions.

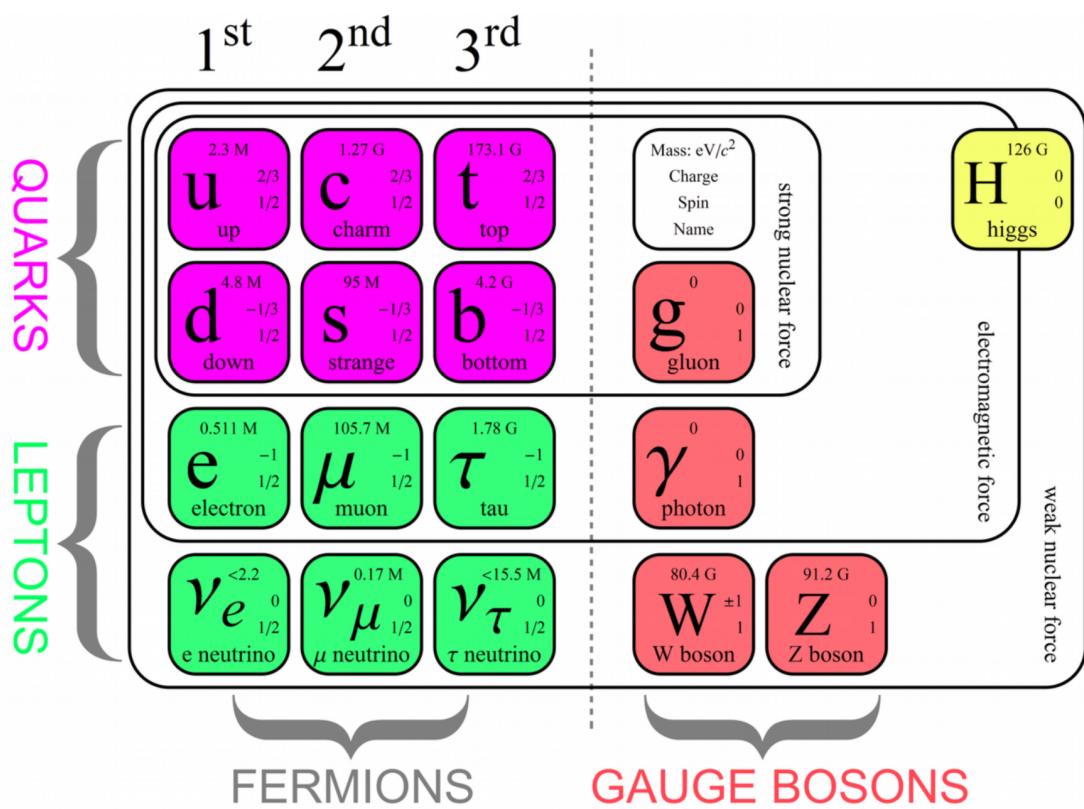


Figure 2-1. The fundamental particles of the Standard Model.

2.2.1 Bosons: Use the Force

Ever wonder how two electrons "know" that they are near each other and that they should repel? Figure 2-2 shows a Feynman diagram of two electrons "communicating" with each other by means of an intermediate photon. I like to think of it as the two electrons "playing catch" with the photon. The first electron recoils from the throw and then the second electron recoils from catching the photon. The photon carries some momentum away from the first electron and brings it to the second one, therefore making it look as if the two electrons are repelling one another! The photon isn't *real* of course - it is said to be a *virtual* photon. Now we see why bosons are called the force carriers.

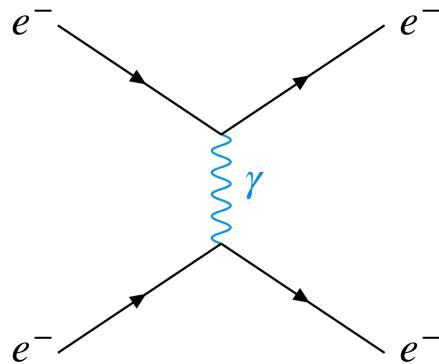


Figure 2-2. A Feynman diagram showing electron-electron scattering, also known as Møller scattering.

A diagram such as the one above is a Feynman diagram and it gives us a wonderfully simple way to visualize particle physics processes. It's not *actually* what happens between the particles, but it is a good starting point. Each diagram is actually a single, scalar number - a complicated QFT integral that tells you how likely a process is to happen. Another benefit to Feynman Diagrams is that they are kind of like tinker toys, in that you can string them together in novel ways to predict real-world processes. Quantum Electrodynamics (QED), one of the theories that make up the SM, mathematically describes and predicts the electrons mediating such a photon between them. It's not just limited to electrons mediating photons, however. QED and other QFTs can predict to astounding accuracy how likely a process is to happen, between whichever particles and fields. You just need to know their properties first.

We have now met the first force carrier: the photon. It is a massless particle and is the mediator of the EM force. Photons only interact with particles that carry *electric charge*. Depending on what kind of charge a particle carries, determines with which bosons it may interact and via which forces. Speaking of forces, the four fundamental forces found in nature, along with their decreasing, relative strength are:

1. strong force (1)
2. EM force (10^{-1})
3. weak force (10^{-6})
4. gravitational force (10^{-40})

If the photon is the mediator of the EM force, then what mediates the other forces?

The mediators of the strong force are the **8 gluons**. Similar to the photon, they are also massless, but that's about all they have in common. Gluons are trapped inside of protons, neutrons, and other hadronic matter. They are responsible for “glueing” nuclei to together, hence their name. Just as photons can only interact with particles that have electric charge, gluons can only interact with particles that have *color* charge. Interestingly, gluons themselves carry color charge which they mediate back and forth between quarks (fermions discussed below). This is quite different from the photon which itself does not carry electric charge. There are three kinds of color charges: red, green, and blue. Every gluon carries two color charges: one kind of color and an anticolor: antired, antigreen, or antiblue.

There are three bosons which mediate the weak force: the Z, W^+ , and W^- . They are extraordinarily massive particles, weighing in at 91.2 GeV for the Z and 80.4 GeV for both kinds of W bosons. That means the W bosons weigh more than an iron atom! These bosons interact with any particle that carries “weak hypercharge”. The weak force has plagued physicists for nearly a century until only recently. Particles which decay via the weak force live an astonishingly long time. Take for example the neutral pion (π^0). It decays very quickly, via the EM force, into

two photons on the order of 10^{-18} s. Now take the charged Kaon K^+ . This particle decays into three charged pions, but takes on average 10^{-8} s to do so. Over 10 orders of magnitude different from the pion decay. This is because the charged kaon decays via the weak force.

The last boson not yet mentioned is the scalar Higgs boson, which is introduced in Section 2.3 below.

2.2.2 Fermions: Each One Matters

There are 12 kinds of fermions - the matter particles of the Universe. They comprise all the “stuff” that we see and feel. All fermions have half-integer spin, typically a value of 1/2. The fermions can be split into two groups depending on if they interact with the strong force (quarks) or not (leptons). Let’s consider the leptons first.

Leptons: We already introduced one lepton earlier: the electron. Looking again at the “particular table”, the electron has a heavier brother, the muon, which is 200 times heavier than the electron. Then there is an even heavier sibling: the tauon. All three of these leptons have the familiar -1 charge which allows them to interact via the EM force and exchange photons with other electrically charged particles.

The charged leptons also carry weak hypercharge, which allows them to interact via the weak force. If a charged lepton interacts with a W^\pm boson, it can transform into its corresponding “partner” - the other member of the $SU(2)$ isospin doublet: the neutrino. These fickle particles are neutral and *only* interact via the weak force (well, and maybe gravity). They are very difficult to detect.

Quarks: The six quarks are the fermions which interact with gluons. They have *quarky* names like: up, down, charm, strange, top, bottom. These are called the six “flavors” of quarks. The top quark is an absolutely massive particle, reaching the top of the mass scale of any particle at 173 GeV - about as heavy as a tungsten atom.

Quarks are electrically charged particles, but they have fractional charge. Each quark in the top row of Fig. 2-1 has +2/3 electric charge and the bottom row has -1/3. That’s why when you combine two up quarks with a down to form a proton quark, the combination of electric charge

yields +1.

Just as the leptons carried weak hypercharge and could interact via the weak force, so too can quarks. The W^\pm bosons can change one flavor of quark into another. The Z boson only affects the spin, momentum, and energy of the particle with which it interacts.

In addition to electric charge, quarks also carry one kind of color charge, either red, green, or blue. It is this color charge which allows them to interact with gluons via the strong force. This is an artifact of being gauge bosons of the $SU(3)$ symmetry group. They combine in different ways to form at least two types of hadrons. The first type is baryons, like protons, neutrons, lambdas (anything that is qqq) and the second type is mesons, like pions, kaons, etas (anything of some form like $q\bar{q}$). For some reason which is not completely understood, only colorless bound states form in nature. Just as a ‘+’ charge would negate a ‘-’ charge, so too would the ‘red’ color charge negate ‘antired’ (as in the case of an observable meson) or even combining red, green, and blue (as in the case of a baryon) would yield a colorless bound state.

Antiparticles: It should be noted that almost every *particle* has a corresponding *antiparticle*, whose charges (e.g., color charge, electric charge) are all opposite the original particle’s charges. Accounting for leptons, quarks, bosons, bound states of quarks, and now antiparticles, it is easy to see why sometimes particle physics is referred to as a “zoo”!

2.3 Electroweak Symmetry Breaking and the Higgs Boson

At large energy scales, like those found during the Big Bang or those produced in the energetic proton-proton collisions of the Large Hadron Collider, the electromagnetic (EM) force and weak nuclear force are one and the same: they are unified. However, at lower temperatures (like room temperature for example), we know that the weak force is very different from the EM force. The former mediates decays of radioactive substances, whereas the latter mediates the excitation of electrons in an atom. So what is responsible for the separation of these two forces, this so-called *electroweak symmetry breaking*?

Upon writing down the equations of motion from the SM Lagrangian (easier said than done), one discovers that all the particles mentioned earlier should have *no mass*. Well that’s a

problem because most particles in nature definitely have mass, like the quarks, leptons, W^\pm , and Z bosons. By introducing a complex SU(2) doublet of scalar fields into the SM Lagrangian, in such a way that it leaves the Lagrangian invariant, then all peace can be restored. This scalar field turns out to be the Higgs field, and its excitations are Higgs bosons. Doing so reveals the particle which should have mass. The process of introducing a Higgs field and breaking the electroweak symmetry is called the **Higgs Mechanism**.

Each particle interacts with the Higgs field with a different strength: in fact, a particle's coupling strength to the Higgs field is exactly its mass! The more the particle interacts with the Higgs field, the more mass it gains. Excitations, or quanta, of the Higgs field are Higgs bosons and are a direct consequence of introducing a Higgs scalar field into the SM Lagrangian to allow particles to have mass.

2.4 The Standard Model Doesn't Explain Everything

The SM has only mathematically accommodated the strong, EM, and weak forces. One problem however is that the SM can't predict the mass of the Higgs boson... or the mass of *any* particle for that matter. That's not the only thing the SM has trouble doing. For example, the SM can't...

- ...incorporate gravity into its mathematical framework.
- ...explain why most of the Universe is made of matter and very little antimatter.
- ...predict the existence of dark matter - but we know it's there from observation.
- ...explain why there should be exactly three generations of fermions.

So we see that the SM isn't the ultimate Theory of Everything, but it does a pretty good job. How can we test the SM and try to break it or confirm it? There are at least two routes to choose from: A patient route and an impatient route. The patient route requires us to wait until our particles of interest maybe come from outer space or, if we produce it in the lab, wait for it to decay into other particles. This could take a VERY long time, (possibly way longer than the age of the Universe - if a proton even decays at all!), or it could take as short as a billionth of a billionth

of a millionth of a second, like in the case of a Z boson. It's not the most reliable method, it is difficult to control, and it requires a lot of patience.

Instead, let's be impatient: let's smash particles together and convert their energies into new kinds of matter. If we use hadrons, which are made of smaller parts like, quarks and gluons (let's call them “partons”) then we will have many more interactions and a lot more fun (Fig. 2-3). We are going to need a lot of energy, so we should make a large collider. Let's make a Large Hadron Collider!

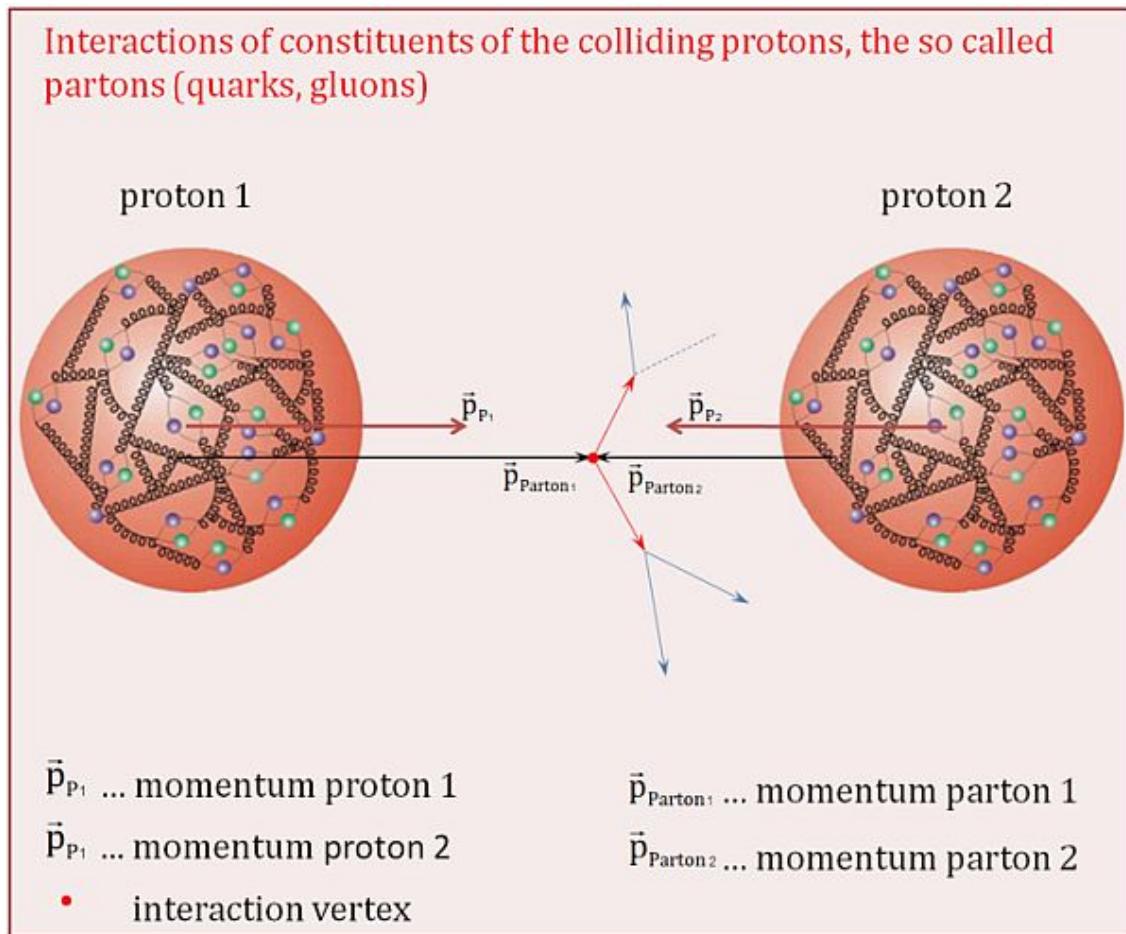


Figure 2-3. Two protons can be smashed together at very high energies to have their partons interact and convert the high energies into new kinds of matter.

- strong force (1)
 - EM force (10^{-1})
 - weak force (10^{-6})

- gravitational force (10^{-40})

Speaking of forces, the four fundamental forces found in nature, along with their decreasing, relative strength are:

2.4.1 Electroweak Interaction

Particles can interact with one another at long range. For example, an electron can emit a photon which can travel a

Electromagnetic force and weak force were unified Electroweak symmetry. This introduced 4 electroweak bosons: the photon, Z, W+, and W-. Then electroweak symmetry breaking happened. This made the photon massless and the Z, W+, and W- as (very) massive.

2.4.2 Strong Interaction

2.4.3 Higgs Mechanism

Force field that fills space in the whole Universe but has no source or direction. The field has the same field at every point. It's a scalar field with spin 0. It does have electroweak charge.

Current problems with the SM. No Gravity. Can't combine quantum mechanics and gravity. No neutrino masses. Are they Dirac or Majorana particles? Higgs field parameters appear highly fine-tuned. Does not explain dark matter or dark energy.

$N_{\text{obs}} \text{ events} = xs * \text{eff} * L \text{ int cross section}$ is specific to the process efficiency is ideally unity.

2.5 Higgs Boson Production at the LHC

CHAPTER 3

THE LARGE HADRON COLLIDER

Located on the border between France and Switzerland, sandwiched between the beautiful Jura mountains to the west and the sprawling city of Geneva (Genève) to the east, is CERN: the European Organization for Nuclear Research (Conseil Européan pour la Recherche Nucléaire). CERN is an international collaboration of more than 23 member states and its “family” is steadily growing. This collaboration is responsible for the construction and commissioning of the world’s largest and most powerful particle accelerator: the Large Hadron Collider (LHC).

The LHC collides particles using the brightest beams and highest energies that humans have ever made. Bright beams meaning highest luminosity.

Higgs boson produced every 1 billion collisions.

Beginning in 2026 (FIXME), the LHC will undergo a “Phase 2” upgrade and become the High-Luminosity LHC. This upgrade will increase the collider’s luminosity by 10 fold (FIXME) and is predicted to deliver SO much data 3000 fb?.

CHAPTER 4

THE CMS DETECTOR



Figure 4-1. Life-size poster of the CMS detector, taken during CERN Open Days 2019 in the SX5 warehouse where parts of CMS were assembled.

Weighing in at 14,000 tonnes, standing 5 stories tall (15 m), and reaching 29 m long, the Compact Muon Solenoid (CMS) experiment is one of two general-purpose particle detectors at the LHC (Fig. 4-1). CMS is situated approximately 100 m under the earth at the fifth collision point (Point 5) along the LHC (Fig. 4-2). In 2012, both CMS and its competing experiment, ATLAS, independently discovered the Higgs boson.

As discussed in Section (TODO: REF), the LHC collides bunches of protons every 25 ns to produce thousands of new particles which then travel away from the interaction point. CMS is built around the interaction point in a series of cylindrical subdetectors for nearly hermetic coverage so that most of the particles must travel through CMS. The detector sports a solenoid, after which CMS was named, which generates a 3.8 T uniform magnetic field that points longitudinally down the central axis of CMS. This strong magnetic field applies a Lorentz force on the outgoing charged particles, causing them to follow helical, momentum-dependent trajectories. These curved tracks are then better separated from one another which assists in particle identification. Neutral particles experience no Lorentz force and thus travel in straight lines.

The subdetectors measure the properties of the outgoing particles and carefully filter them out in a clever way (Fig. 4-3). Particles interact with the subdetectors, leaving so called “hits”

where they passed through. Hits are reconstructed into tracks. From the track curvature, deduce charge and momentum of the particles. Depending on which subdetector (or combination of subdetectors) was hit by the outgoing particles, the type of particle can be deduced. A few

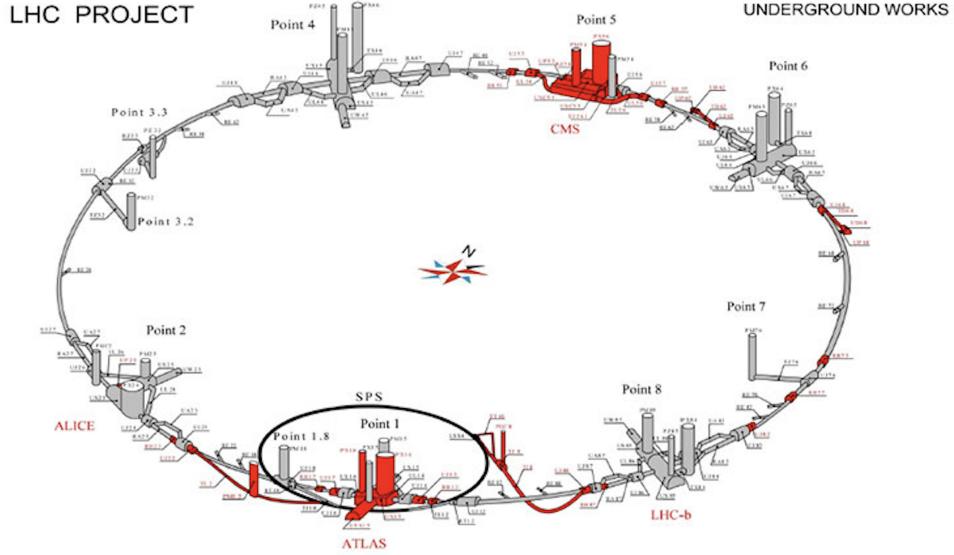


Figure 4-2. Points 1 through 8 along the LHC. Collisions occur at Points 1 (ATLAS), 2 (ALICE), 5 (CMS), and 8 (LHCb), whereas the remaining points are used for LHC beam maintenance and testing.

example particles and their associated tracks are shown in Fig. 4-4.

Before discussing each subdetector in the following sections, it is useful to define the coordinate system used in CMS: a typical, right-handed, three-dimensional Cartesian coordinate system (x, y, z) is used, whose center $(0, 0, 0)$ is placed at the nominal pp collision point within CMS. The x -axis points towards the center of the LHC, the y -axis points vertically upward, and the z -axis points westward towards the Jura mountains, tangential to the beam direction. Since CMS covers almost the entire spherical 4π steradians around the interaction point, it is convenient to use spherical coordinates (r, ϕ, θ) , in which r measures the radial distance in the x - y plane, ϕ measures the azimuthal angle in the x - y plane as measured from the x -axis, and θ measures the polar angle as measured from the z -axis. When dealing with ultra-relativistic particles like those produced at the LHC, special relativistic effects like length contraction must be taken into account and so the coordinate θ becomes frame-dependent. It is thus helpful to convert θ to the

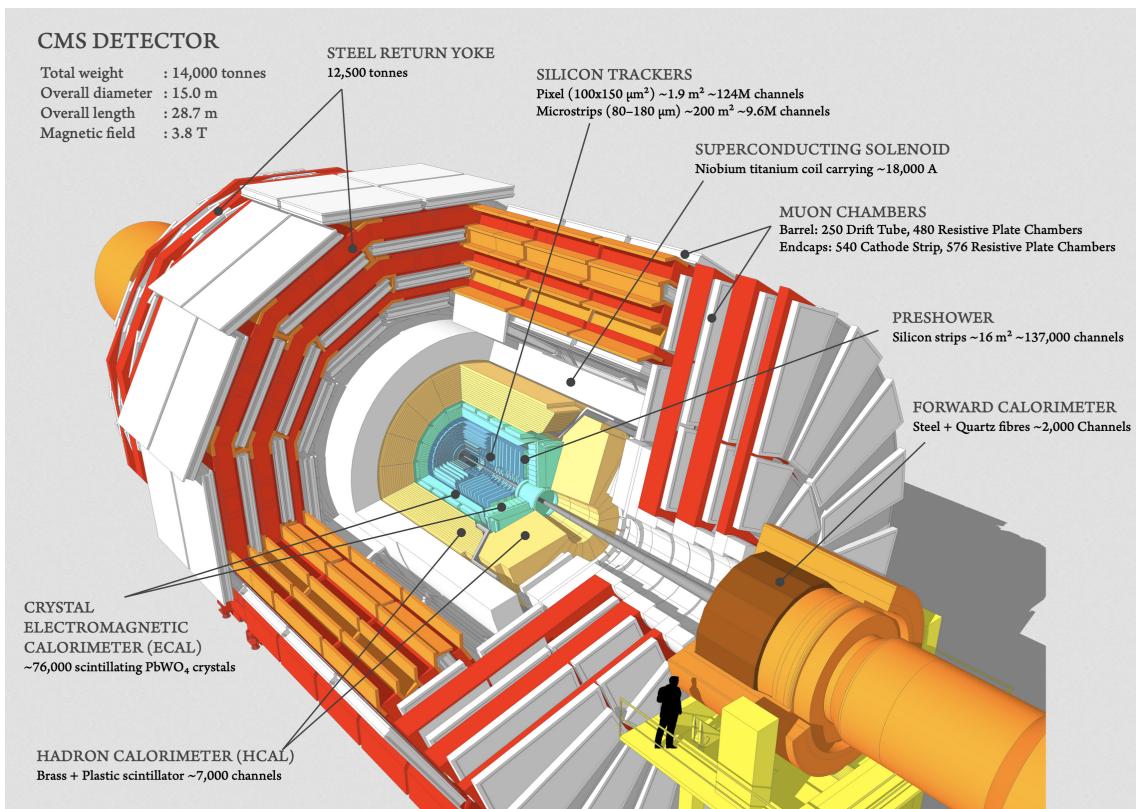


Figure 4-3. Cut out of the CMS detector showing its various subdetector components.

Lorentz-invariant quantity called pseudorapidity (η), defined as $\eta = -\ln[\tan(\theta/2)]$.

4.1 The Silicon Tracker

At the heart of CMS is one of the world’s largest silicon detectors: the silicon tracker. The main goal of the silicon tracker is not to capture outgoing particles but to very precisely measure the hits from the charged particles as they pass through it. The tracker also assists in vertex identification, differentiating between primary and secondary vertices, the latter of which often comes from B meson decays. When multiple pp collisions occur within the same BX (pile up), the tracker distinguishes between proton collisions with a resolution of about $100 \mu\text{m}$ longitudinally and $50 \mu\text{m}$ transverse to the beam pipe. This is crucial to resolve which outgoing particles came from which pp vertex.

The tracker consists of two types of pure silicon detectors: the pixel detector and the strip detector, each of which is described in detail below.

4.1.1 The Pixel Detector

The innermost part of the silicon tracker is the pixel detector, which is the closest subdetector to the interaction point. The pixel detector is composed of 66 million silicon “pixels”, as shown in Fig. 4-5 (Left, pink). A single pixel is $100 \mu\text{m} \times 150 \mu\text{m}$ and, collectively, they cover a sensitive area of 1.9 m^2 . Because it sits only 8 cm away from the beam pipe, the pixel detector receives the highest particle flux than any other subdetector: around 10 million particles per cm^2 per second.

The pixel detector is made of three cylindrical layers and two endcaps that surround the beam pipe. In total, the pixel detector has around 6,000 connections (channels) per cm^2 .

After the LHC Run 1 was completed, the accelerator received luminosity upgrades during the 2013-2014 long shutdown period. To handle these higher luminosities, the pixel detector was replaced by the CMS Phase-1 pixel detector during the LHC technical stop in 2016-2017. The upgrades outfitted the detector with four barrel layers and three endcap disks per side, which allowed for particle detection up to $|\eta| < 2.5$. The overall mass of the pixel detector decreased and granted the detector with better tracking capability.

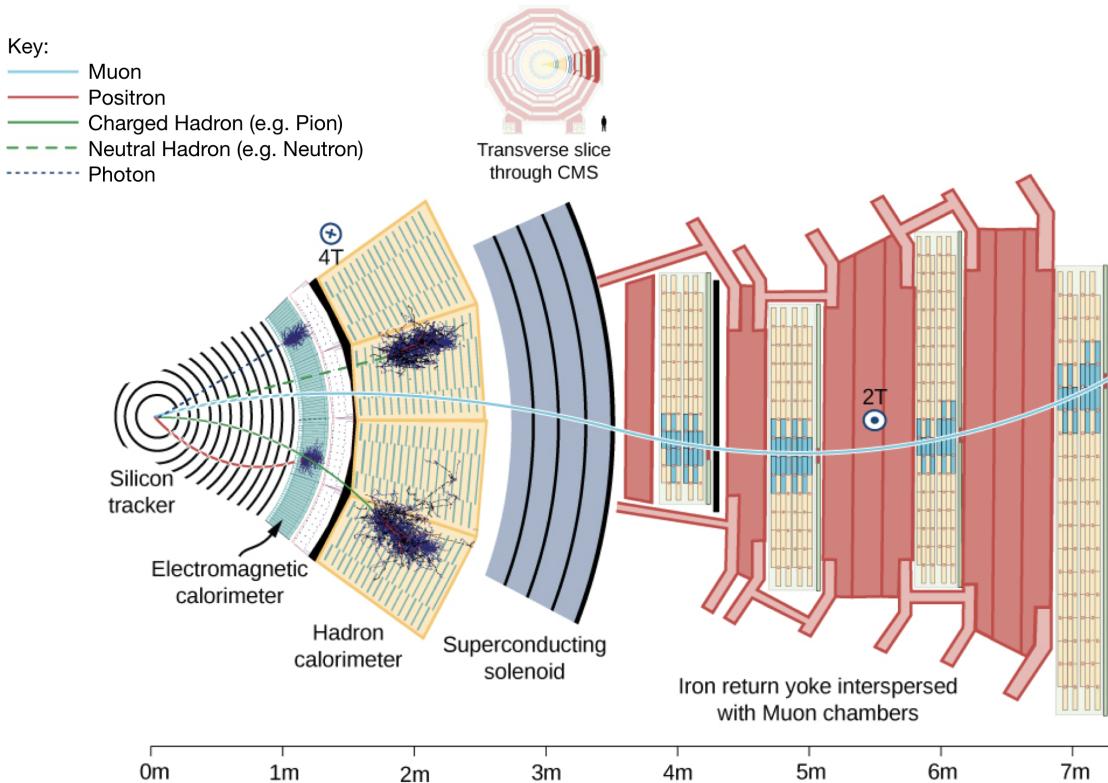


Figure 4-4. A transverse view of CMS showing the “filtration process” as different particles pass through different subdetectors. A positron (solid red line) curves due to the presence of the magnetic field and gets stopped in the ECAL, creating an EM shower. A photon (blue dashed line) does not get detected at all by the Silicon Tracker, since it has no electric charge. It continues through to the ECAL and makes a shower here, like the positron. Charged hadrons (solid green line) will show curved tracks from the Silicon Tracker, may leave some trace in the ECAL, but primarily get stopped by the HCAL creating hadronic showers. Neutral hadrons (dashed green line) do not interact with the tracker, and only undergo EM showers a little in the ECAL, but show most energy deposits in the HCAL. Muons (solid blue line) are detected by the Silicon Tracker and then mostly pass through the other subdetectors without interacting until they finally reach the Muon System. Using the Lorentz force law and knowing which direction the magnetic field is pointing, one can deduce the sign of the charge of the particle. Based on the radius of curvature from the trajectory, one can then calculate the momentum and energy of the particle.

4.1.2 The Strip Detector

The outer part of the silicon tracker is called the strip detector, which has 10 million detector strips spread across 10 cylindrical layers. The first 4 layers belong to the tracker inner barrel (TIB) and the remaining 6 layers belong to the tracker outer barrel (TOB), Fig. 4-5 (Left, green and blue, respectively). Both the TIB and TOB have two endcaps associated with them, the TID and TEC, respectively. Accounting for all of its components, the strip detector is sensitive to 200 m^2 .

Fig. 4-6 gives a clearly-labelled transverse illustration of the pixel and strip detectors.

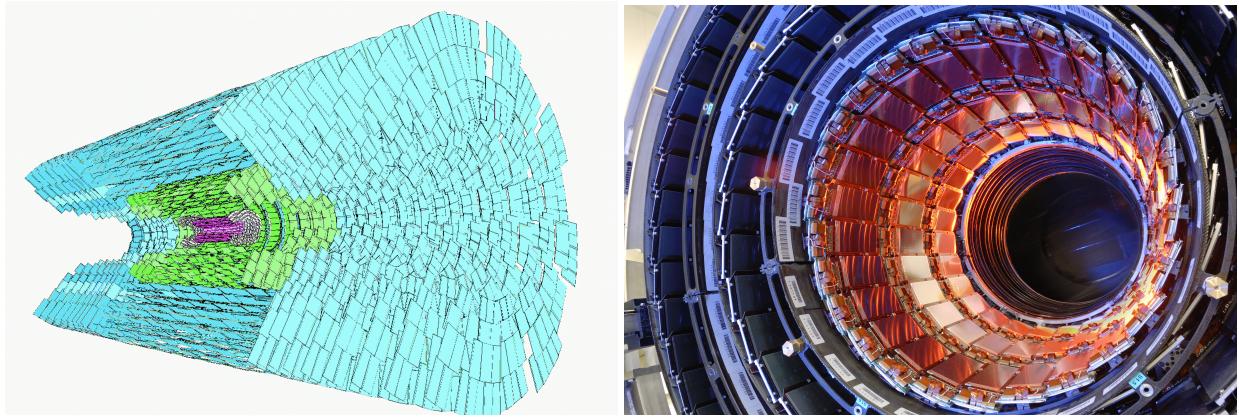


Figure 4-5. (Left) A simulation of the silicon tracker, showing the 3 cylindrical layers of the pixel detector (pink), 4 layers of the TIB (green), and the 6 layers of the TOB (blue) of the strip detector. The endcap components are also shown. (Right) A picture of the real silicon tracker at the center of CMS.

4.2 The Calorimeters

4.2.1 Electromagnetic Calorimeter

Overview: Particles that pass through the silicon tracker 4.1 encounter the electromagnetic calorimeter (ECAL). Those particles which interact electromagnetically but not strongly, mostly photons and electrons, are typically absorbed by the ECAL. The particle's energy is then transferred to the ECAL in the form of an electromagnetic (EM) shower. The size and shape of the EM shower provide information about the particle's energy and trajectory. Since the Higgs boson can decay into two photons, it was essential that the ECAL was able to detect this decay mode.

Design: The ECAL is a hermetic, cylindrical, homogeneous sub-detector that consists of a barrel (EB), two endcaps (EE), and a preshower detector in front of each endcap (Figure 4-7,

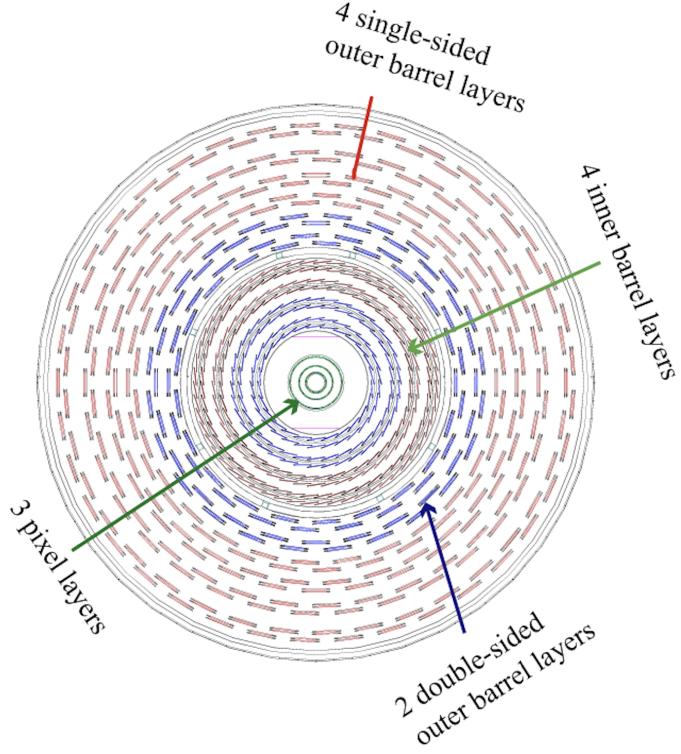


Figure 4-6. A transverse view of the silicon pixel and strip detectors, explicitly labelling the different layers involved.

Left). The EB covers $|\eta| < 1.479$ while the EE covers $1.479 < |\eta| < 3.0$. The entire subdetector is composed of transparent lead tungstate (PbWO_4) crystals that point axially towards the center of CMS (the interaction point). The transparent crystals, one of which is shown in Fig. 4-8 (Left), have a high density (8.28 g/cm^3) which provides the ECAL with radiation resistance and a short radiation length ($X_0 = 0.89 \text{ cm}$). Because so many crystals are used (61,200 crystals in the EB and 7,324 in the EE), the ECAL has excellent energy resolution and fine granularity. Each endcap is composed of two “Dee”s, one of which is shown in Figure 4-7 (Right). A single Dee carries 3,662 crystals. Crystals in the barrel are tapered, having front face dimensions $2.2 \times 2.2 \text{ cm}^2$, back face dimensions $2.6 \times 2.6 \text{ cm}^2$, and are 23.0 cm long ($25.8 X_0$). Crystals in the endcaps are also tapered, with front face dimensions $2.862 \times 2.862 \text{ cm}^2$, back face dimensions $3.0 \times 3.0 \text{ cm}^2$, and are 22.0 cm long ($24.7 X_0$). This gives a single crystal from the barrel a volume of approximately 132.5 cm^3 (mL), about the volume of a small cup of coffee, yet it weighs 1.5 Kg. (REF:PDG).

REF:PDG Particle Data Group collaboration, S. Eidelman et al., *Review of particle physics*,

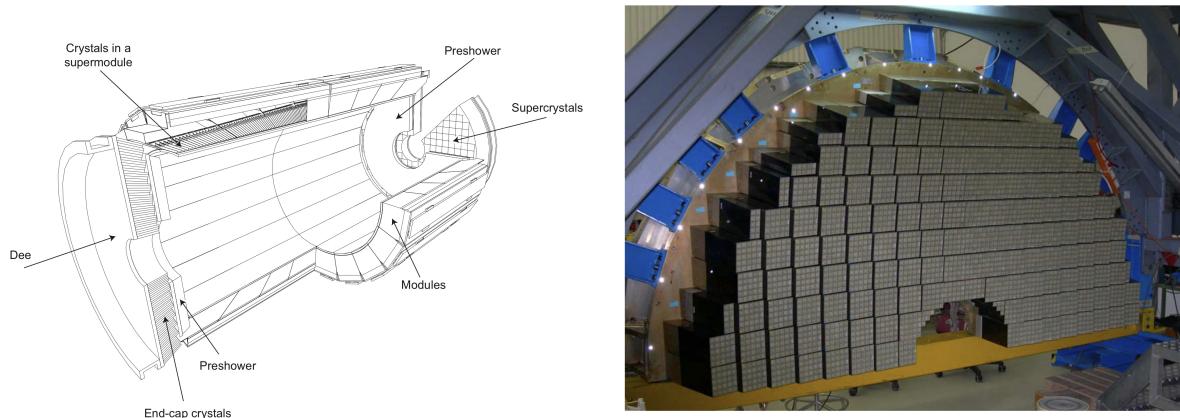


Figure 4-7. (Left) Cross sectional view of the electromagnetic calorimeter of CMS. (Right) One of the Dees which comprise the EE. Each square of 5×5 crystals constitutes a “super-crystal”. Figure taken from Ref. [1].

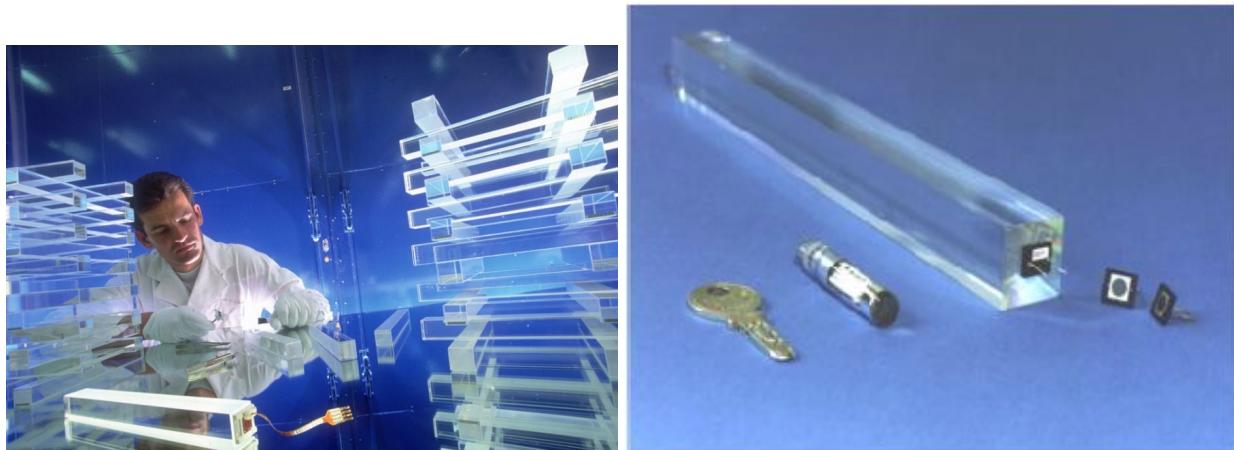


Figure 4-8. (Left) ECAL crystals made from PbWO₄ are grown in a lab. (Right) Although made mostly of metal, ECAL crystals are transparent and have a photomultiplier detector attached at the end.

Physics: When electrons or photons pass through the ECAL, they create an EM shower. Electrons radiate more photons as they accelerate around PbWO₄ nuclei, in a process called *bremsstrahlung*. Meanwhile, near the presence of a nucleus, high-energy photons pair produce into e⁺e⁻. This cycle of electron and photon production disperses the initial particle energy into a spray of lower- and lower-energy particles; an EM shower (REFERENCE EM SHOWER FIG).

SHOW PICTURE OF EM SHOWER.

The ECAL crystals then scintillate (emits photons) in proportion to the amount of energy deposited by the interacting particle. The scintillator photons are detected by avalanche photodiodes on the back of each barrel crystal or by vacuum phototriodes in the endcap crystals (Fig. 4-8, Right). Conveniently after 1 bunch crossing (25 ns), Approximately 80% of the scintillated light is emitted.

An energy deposit in the ECAL could come from either an electron or a photon. In order to tell the difference, information from the silicon tracker is used. Charged particles, like electrons, will leave hits in the tracker and follow a curved path, whereas photons are electrically neutral and thus will not show any signs within the silicon tracker. So long as the tracker and ECAL communicate effectively with each other, then they help distinguish between electrons and photons. Charged hadrons interact only minimally with the ECAL, instead continuing on to the Hadron Calorimeter. Neutral hadrons can be detected by the ECAL preshower near the ECAL endcaps which helps distinguish a single photon from π^0 mesons as they decay into two photons with a narrow opening angle, making it look as if the two photons are a single photon. The preshower detector allows CMS to distinguish between low-energy diphoton pairs and single high-energy photons.

NEED SMOOTH TRANSITION INTO HCAL. What about those hadrons? They got through the ECAL... To detect hadrons effectively, we need a Hadron Calorimeter.

4.2.2 Hadron Calorimeter

Overview: The particles that survive the ECAL, typically only muons and hadrons, then enter the hadron calorimeter (HCAL). Its primary purpose is to absorb the hadronic matter emerging from the interaction point and measure the corresponding jet energies. The absorbed jets cause the HCAL to scintillate photons which are then converted into electrical signals. These signals are used to deduce the original jet energies and any missing transverse energy (E_T^{miss}) from the event.

Design: Dissimilar to the ECAL (subsec. 4.2.1) in material composition but similar in shape, the HCAL is a brass cylindrical scintillator. Although it has a barrel (HB) and two endcaps

(HE), it has two more detectors than the ECAL: the outer calorimeter (HO) and the forward calorimeter (HF). The HB spans the pseudorapidity range $|\eta| < 1.3$, the HE spans $1.3 < |\eta| < 3$, and the HF spans $3 < |\eta| < 5.2$, as shown in Figure 4-9. With a thickness of over 1 m, the HB is sandwiched between the barrels of the ECAL and the solenoid (subsec. 4.3) at radial values $r = 1.77$ m and $r = 2.95$ m, respectively. Because the HB and HE are located within the solenoid's strong magnetic field of 3.8 T, they both were both constructed out of a non-magnetic absorber called *C26000 cartridge brass*. This absorber has a density of 8.53 g/cm^3 and an interaction length (λ_I) of 16.42 cm. The thickness of the HB increases as $1/\sin\theta$ so that at $|\eta| = 0(1.3)$ the absorber thickness is $5.82(10.6) \lambda_I$. The HB is composed of two half-barrels, where each half-barrel is built from 18 identical azimuthal wedges and each wedge spans 20° . Each wedge is divided into four ϕ segments so that a single ϕ segment spans $\Delta\phi = 0.087$.

Since the volume available to the HCAL is so limited, and in order to stop any particles that might traverse the entire HCAL and solenoid, the HO (the “tail catcher”) is situated outside the barrel of the solenoid. The HF is located 11.2 m from the interaction point.

All tiles within a single ϕ segment are grouped together into a single tray unit. The scintillator is also segmented into 16 η sectors, the first(last) of which is located at $|\eta| = 0(1.3)$. This way each tile covers $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$. Each layer has 108 trays.

Physics: Since hadrons are the only particles to interact via the strong force, the HCAL is designed to have a high nuclear density. This ensures ample opportunity for hadrons to radiate gluons and convert with the Similal to the ECAL, the HCAL will scintillate in proportion to the amount of energy of the captured particle. The incoming hadrons will *hadronize* (i.e., produce a hadronic shower), generating jets of quarks and gluons which are bound in various ways forming protons, neutrons, pions, kaons, etc. Interestingly, the HCAL is made using over a million old, brass shell casings from the Russian Navy back from World War II.

About 34% of the particles produced from LHC pp collisions enter the HE region, so the HE was built to handle high rates (MHz).

The entire HCAL utilizes approximately 70,000 plastic scintillator tiles. The active material

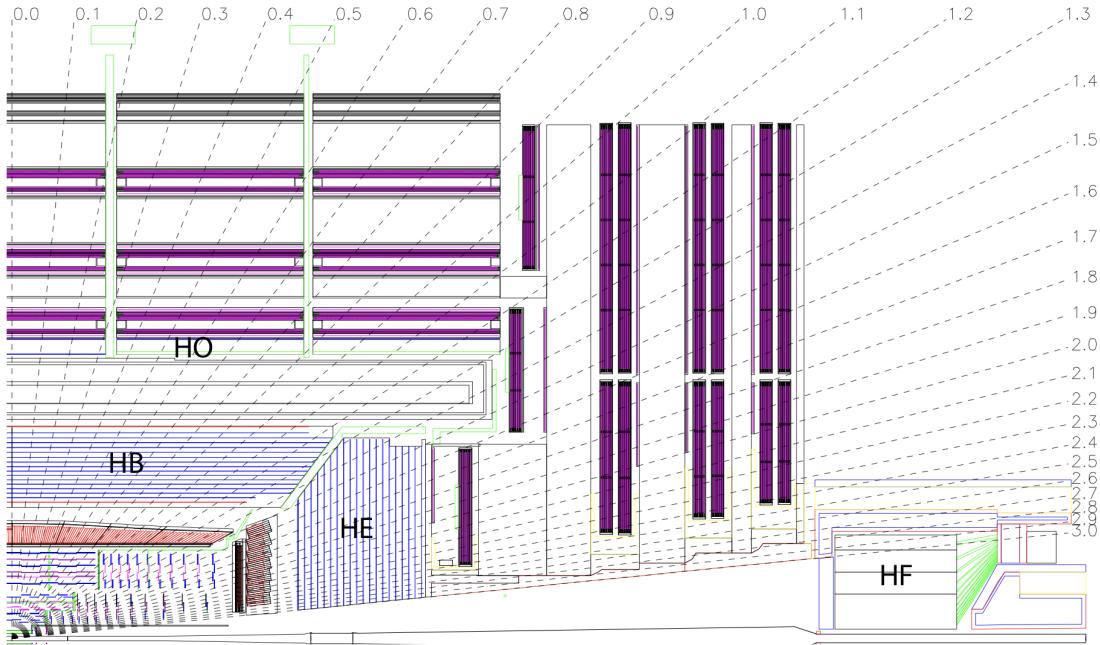


Figure 4-9. A cross-sectional quadrant of CMS showing the locations of the HCAL components: the barrel (HB), outer (HO), endcap (HE), and forward (HF) detectors.

in the HB is 3.7-mm-thick Kuraray SCSN81 plastic scintillator, selected for its radiation hardness and long-term stability. Hadron showers \rightarrow tiles scintillate \rightarrow scintillated photons are collected by 0.94-mm-diameter green double-cladded wavelength shifting (WLS) fibers (Kuraray Y-11), which carry the light to hybrid photodiodes (HPD).

4.3 The Solenoid and the Steel Return Yoke

The Compact Muon *Solenoid* sports one of the world's most energetic solenoids which is paramount to the success of CMS. Particles that exit the HCAL (subsec. 4.2.2) arrive at the cylindrical magnet which is 12.5 m in length, has a bore diameter of 6 m (6.3 m when cold), and generates a uniform 3.8 T magnetic field parallel to the beam line. To produce such a large and uniform magnetic field inside the approximately 360 m^3 volume (Fig. 4-10), an 18,000 amp current travels through the 4-layer, superconducting, NbTi coils. This magnetic field is 100,000 times stronger than Earth field at the surface, storing a massive 2.6 GJ of energy—approximately the kinetic energy of an Airbus A320 in flight. The magnet has such a

large stored-energy-to-cold-mass ratio (11.6 KJ/Kg) that it experiences a physical deformation of 0.15% while energizing the field.

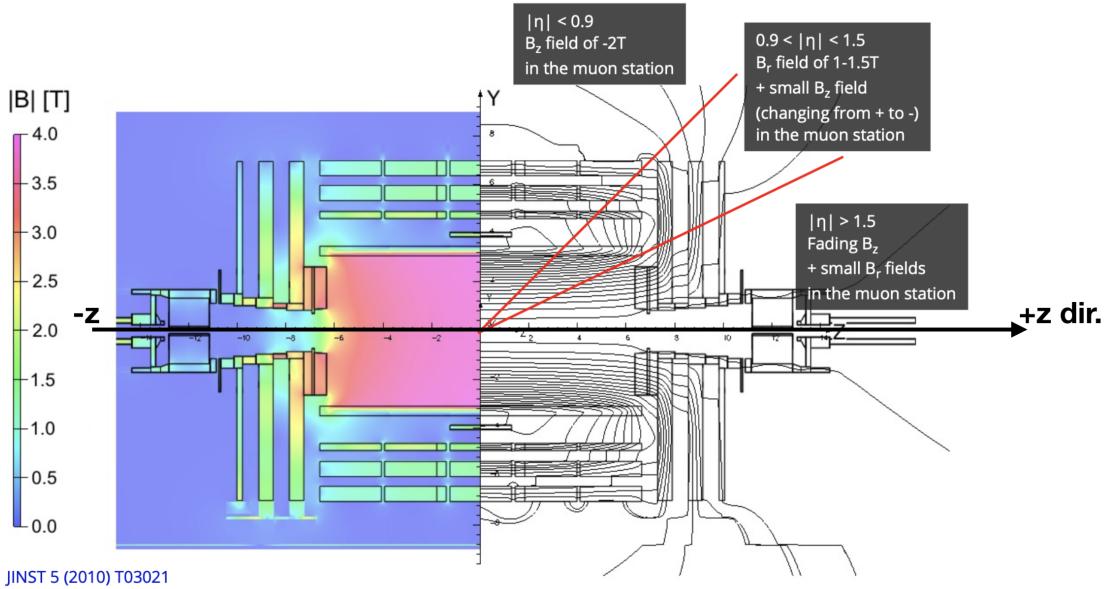


Figure 4-10. A longitudinal cross section of CMS showing the values of the magnetic field over the volume of CMS and various field lines. The magnetic field reaches its maximum of 3.8 T in the center of the detector.

As charged particles travel through any magnetic field, they experience a magnetic (Lorentz) force perpendicular to their direction of travel. The Lorentz force (\vec{F}_B) exerted on a particle with charge q depends on the particle's velocity (\vec{v}) and the strength of the magnetic field (\vec{B}), given by

$$\vec{F}_B = q\vec{v} \times \vec{B}$$

Since the force is necessarily perpendicular to the velocity, the resulting trajectory is a helix. Projecting the helix on the x - y plane (since the magnetic field points in the $+z$ direction) allows the particle tracks to typically be separated from one another. Each track has a corresponding radius of curvature (R) which relates to its transverse momentum (p_T) through

$$p_T = qBR.$$

The relative change in p_T (i.e. the momentum resolution) is given by

$$\frac{\delta p_T}{p_T} \propto \frac{p_T}{BL^2}. \quad (4-1)$$

Steel Return Yoke: Most of the mass of CMS comes from the *steel return yoke* which helps to redirect the magnetic field back on itself. The yoke system constitutes 10,000 tonnes, which is 89% of the total mass of CMS. It is comprised of 5 wheels and 2 endcaps

4.4 The Muon System

Although it is the farthest system from the interaction point, the muon system is one of the most important within the Compact *Muon* Solenoid detector. Of the particles emerging from the interaction point, electrons and photons are absorbed by the ECAL (subsec. 4.2.1) and hadronic matter is absorbed by the HCAL (subsec. 4.2.2). This filtration process leaves only neutrinos and muons to enter the muon system, which is the outermost detector situated past the solenoid (sec. 4.3). As mentioned in , neutrinos are the only weakly-interacting, electrically-neutral SM particles which makes them incredibly difficult to detect directly. In fact, neutrinos interact with normal matter so little that a light-year (9.46 trillion Km) of lead would only stop half of the neutrinos moving through it. Thus, the detection of neutrinos produced from pp collisions is inferred via E_T^{miss} on a per-event basis. Muons, on the other hand, have a mass of 105.7 MeV (relatively heavy for a weakly-interacting, electrically-charged particle) and live a billion-billion times longer than a Higgs boson: the average lifetime of a muon is $\tau_\mu = 2.2 \times 10^{-6}$ s. These properties are what determined the properties of the muon system within CMS to consist of its four main subdetectors, each of which is described in the following subsections:

1. CSC (cathode strip chambers, subsec. 4.4.1),
2. DT (drift tubes, subsec. 4.4.2),
3. RPC (resistive plate chambers, subsec. 4.4.3),
4. GEM (gas electron multiplier, subsec. 4.4.4).

4.4.1 Cathode Strip Chambers

Overview: A cathode strip chamber (CSC) is a multi-wire proportional chamber capable of precisely measuring the position of muons which enter and ionize the gas within. Spatial coordinates are obtained by the collection of electrical signals along the cathode strips (ϕ), anode wires (r), and across multiple CSC layers (z). CSCs are found exclusively on the two endcaps of CMS, with each endcap bearing 270 chambers (Fig 4-11). The chambers are arranged azimuthally around the beam pipe in four disks per endcap allowing for contiguous coverage in ϕ . In total, the CSC system provides an effective detection area of about 5,000 m², has a total gas volume that exceeds > 50 m³, and contains over 400,000 read-out channels.

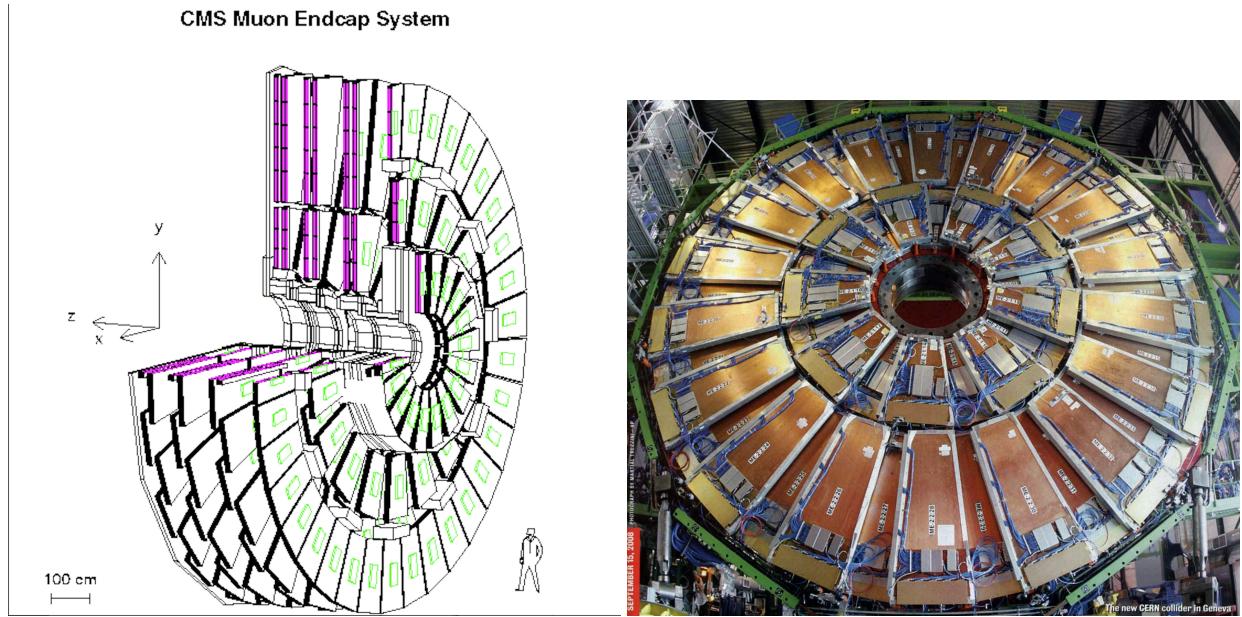


Figure 4-11. (Left) A simulated cut out view of the ME+ endcap, with the coordinate system of CMS in the top-left corner. (Right) The actual ME-2 disk of CMS is shown, revealing its ME-2/1 and ME-2/2 rings of CSCs.

Design: Each CSC is trapezoidal in shape with its narrow end pointed toward the beam pipe (Fig. 4-14, Left). The chambers are arranged in rings and each chamber subtends either 10° or 20° in ϕ , as described in subsubsec. 4.4.1.1. The CSCs cover the pseudorapidity range of $0.9 < |\eta| < 2.4$ (Fig. 4-12). A single CSC is composed of six layers (or *gas gaps*), each of which is filled with a carefully prepared gaseous mixture¹ of Ar:CO₂:CF₄ (Fig. 4-13). The gas mixture

¹The gas mixture ratio of Ar:CO₂:CF₄ was chosen to be 5:4:1 to maximize the lifetime of the CSCs as they endure

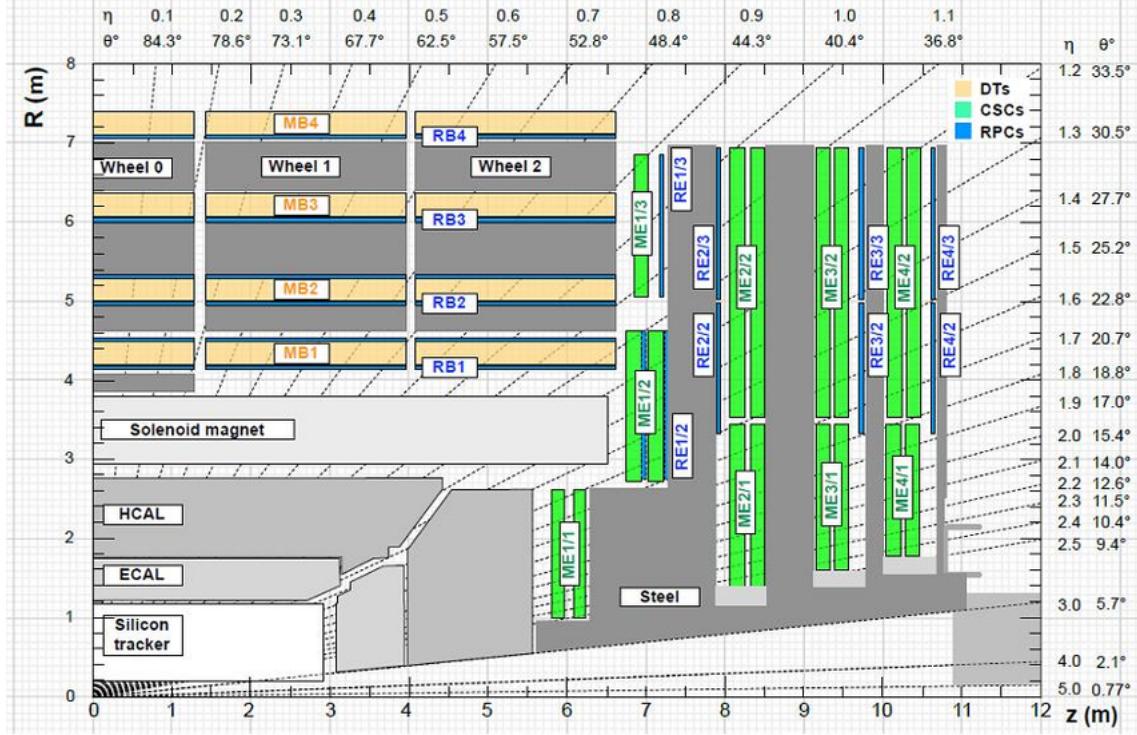


Figure 4-12. Longitudinal cross section of CMS, showing the different pseudorapidity values (η) and also the different subdetector regions.

flows from one layer to the next in a zigzag manner and has a flow rate of ????. Within every layer, the gas mixture surrounds approximately 80 copper strips, each of which spans radially away from the interaction point. A single strip is about 8.4 mm wide at the narrower end of the CSC, about 16 mm at the wider end, and is separated from its neighboring strip by about 0.5 mm. Per layer, the inner gas also surrounds over 1,000 gold-plated tungsten wires, which are oriented azimuthally (so approximately orthogonal to the strips). Each wire is approximately 50 μm in diameter and separated from its neighboring wire by about 3.2 mm. A collection of 16 consecutive wires forms a *wire group*, which is about 5 cm wide and creates a single anode read-out channel. A single wire plane has five independently controlled HV segments (Fig. 4-14, Right). The largest CSCs are 3.4 m long as measured along a strip and 1.5 m wide as measured along a wire.

Physics: As a muon passes through a CSC layer, it has the opportunity to interact with and ionize an Ar atom in the gas mixture. The wires are under high voltage (3,600 V) which causes radiation damage through years of use. The CO₂ is used as a non-flammable quencher to reach even larger electron multiplicities, while the CF₄ helps prevent polymerization along (aging of) the wires.

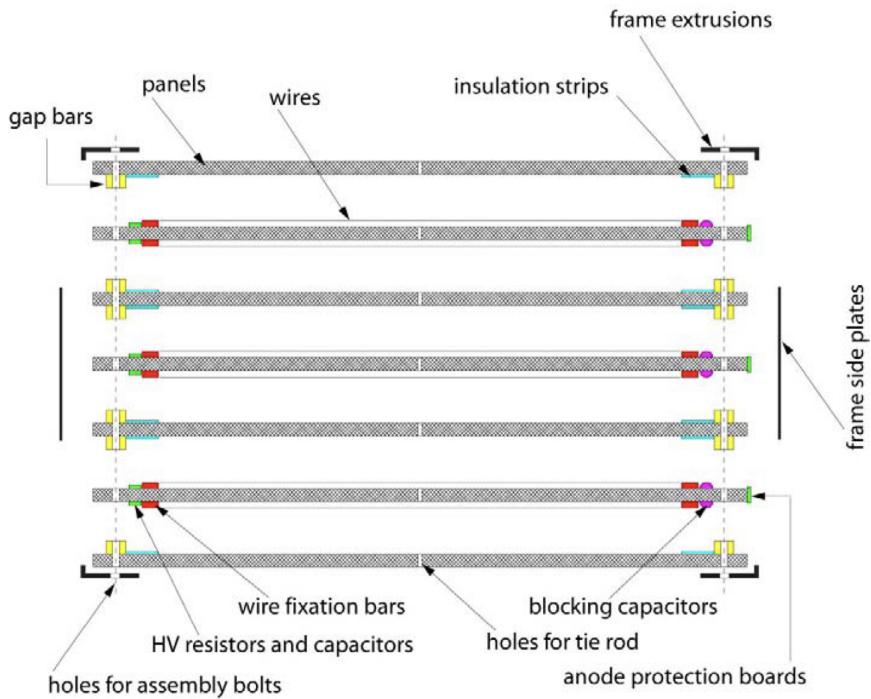


Figure 4-13. Exploded view of the cross section of a CSC showing how the 7 panels come together to form the 6 gas gaps between the panels.

the ionized electron to accelerate towards the positively-charged strips. The accelerating electron collides with and ionizes Ar atoms along its path toward the strip. This liberates even more electrons, thus forming an *electron avalanche* (Fig. 4-15). The total number of ionized electrons per initial electron is referred to as the *multiplicity* (or *gas gain*), which can reach as high as 100,000.

The electron avalanche is collected by a cathode strip and becomes an electrical signal. This signal is processed by the cathode front-end boards (CFEBs). The Ar^+ ions similarly distribute a charge signal onto the negative wires. The cluster of charge that arrives at a strip is more widely spread than the charge which arrives along a wire. Therefore, comparator logic is implemented on the strips to narrow down the precision to the order of $100 \mu\text{m}$ by using half-strip information (Fig. 4-16).

The muon passes through the next 5 layers of the CSC, further ionizing the gas mixture and generating electrical signals along the wires and strips. A signal on a wire provides an

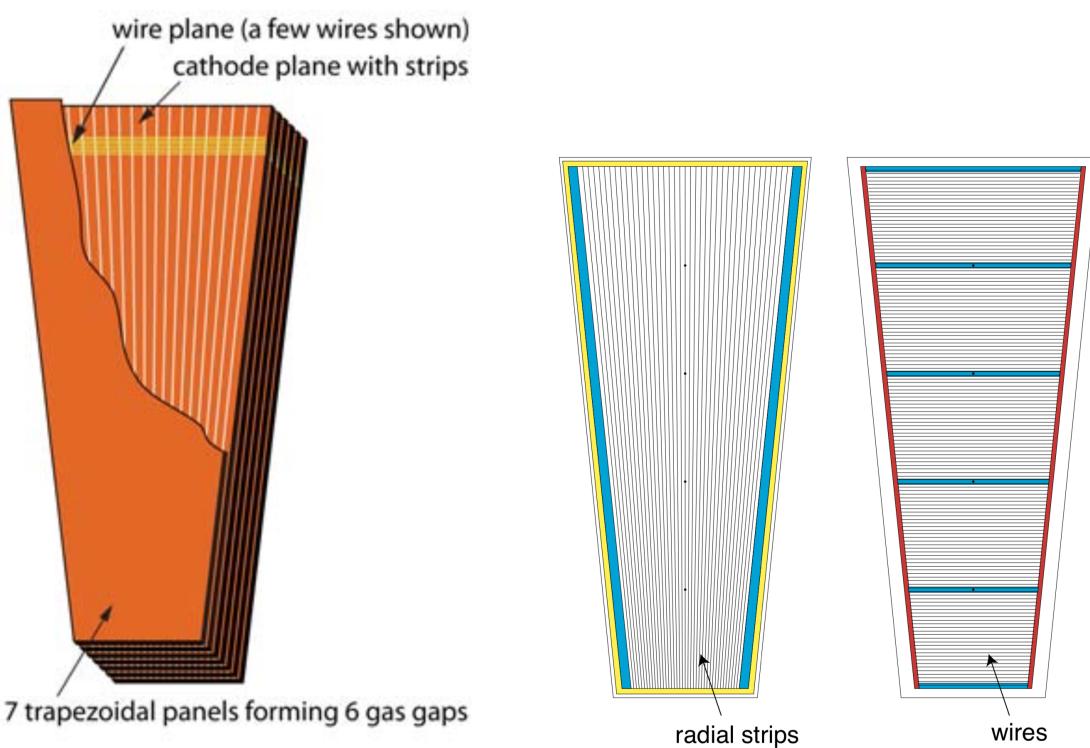


Figure 4-14. (Left) A CSC with its top layer exposed. You can see very thin gold-plated tungsten wires which actually span the entire width of the CSC. Thicker vertical strips run along the length. (Right) More detail showing the radial strips and the horizontal wires. Also shown is the 5 segments of a CSC.

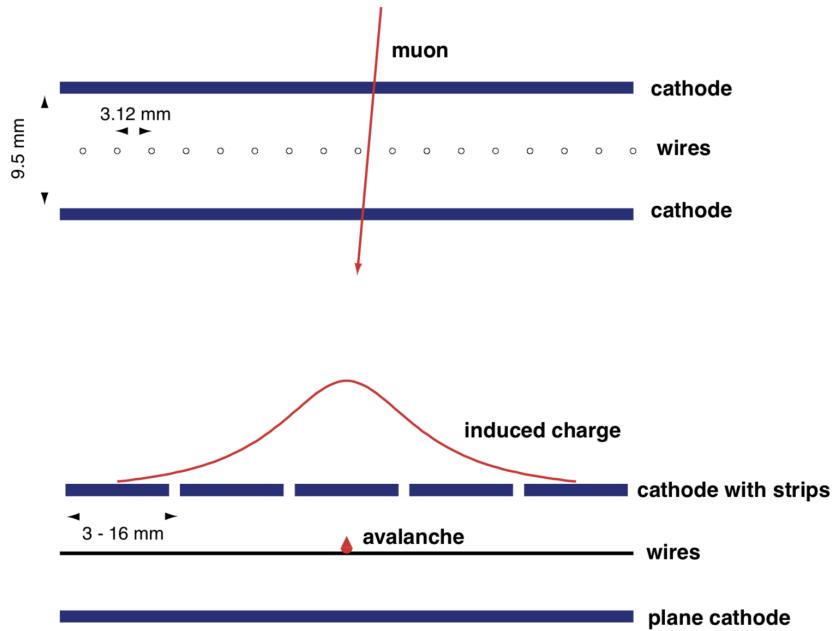


Figure 4-15. A muon passes through one of the gaseous layers of the CSC, ionizing the gas mixture and inducing a charge on the anode wires and cathode strips.

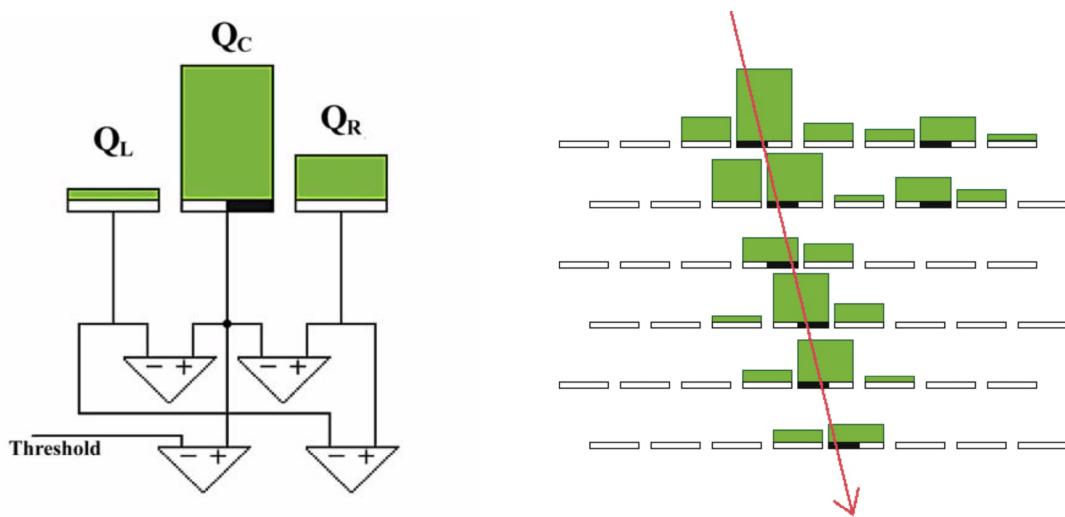


Figure 4-16. (Left) Comparators are used to compare neighboring strip cluster charge to determine on which half-strip the peak charge resided. (Right) A muon passes through all six layers of a CSC inducing charge on various half-strips.

r -coordinate, on a strip provides a ϕ -coordinate, and through the other layers provides a z -coordinate. Taken together, this information helps to reconstruct the three-dimensional trajectory of the muon.

The number of hits recorded by the CSC will determine if an event was significant enough to be worth saving. If so, then its precise positions on the wires and strips will be read out by the Data Acquisition (DAQ) system and be stored for further data analysis. When a CSC is taking live data it can resolve approximately 2 mm in r - ϕ , whereas during offline analysis the resolution improves by a factor of more than 20: the ME1/1 and ME1/2 chambers can resolve distances as small as 75 μm in r - ϕ , while the other chambers can resolve 150 μm . It is worth noting that a CSC can accommodate up to 1 KHz/cm².

Gas gap is 9.5 mm.

4.4.1.1 CSC Numbering Scheme

The two endcaps are labelled as “ME+” and “ME-”, depending on whether they are situated in the $+z$ direction or $-z$. Both endcaps are structurally symmetric, so it is sufficient to discuss only one in detail. The ME+ endcap has four disks: ME+1 is the first disk and the one closest to the interaction point, while ME+4 is the fourth and the farthest away. Within each disk, there are either two or three “rings” of CSCs, as shown in Fig. 4-12 (green). These rings are labelled as ME+ D/R , where D indicates the disk number and R indicates the ring number. For example, ME+2/1 refers to the second disk and the first ring (the ring closest to the beam pipe). All rings contain 36 CSCs, except for ME $\pm X/1$, for $X = 2, 3, 4$, which contain only 18 CSCs. Finally, the CSCs are given one final number to label them on the ring: the CSC that sits along the positive x -axis in the coordinate system of CMS is given the number “01”, e.g. ME+4/2/01. The CSCs are then numbered incrementally following the positive azimuthal direction.

4.4.2 Drift-Tube Chambers

Overview: Functionally similar to but structurally different from a CSC (subsec. 4.4.1), a DT is a collection of gaseous detector cells (Fig. 4-17). A single DT cell has an anode wire and two cathode strips and operates on the same principle as a CSC, providing timing and position

measurements of muons (Fig. 4-18). While CSCs are found only on the endcaps of CMS, drift-tube chambers (DTs) are placed exclusively along the barrel (Fig. 4-19). The DT system is therefore composed of concentric cylindrical stations, with the central axis parallel to the beam pipe. Altogether, there are 250 DTs built inside of, between, and outside of the iron yoke.

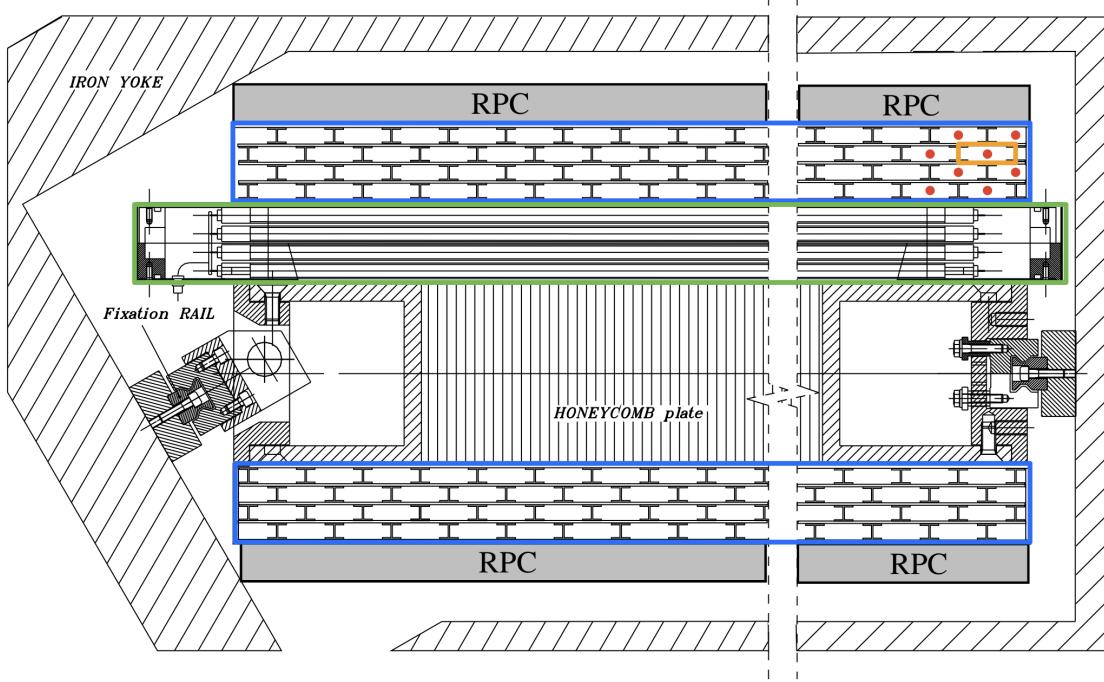


Figure 4-17. A drift-tube chamber with 3 superlayers (SLs) is sandwiching a honeycomb plate. Two SLs are indicated by the blue rectangles, whose cells (orange rectangle) have anode wires (red dots) oriented parallel to the beam pipe. The third SL is indicated by the green rectangle and is staggered orthogonally to the other two SLs.

Design: The first three stations contain 60 DTs each, while the station farthest from the beam pipe contains 70 DTs. The DTs are distributed among 5 wheels, 4 stations, and 12 sectors within the muon barrel (MB) system which uses the following numbering scheme: MB/W/A/S, where *W* is the wheel number (-2 to 2), *A* is the station number (1 to 4), and *S* is the sector number (1 to 12). This accounts for only $5 \times 4 \times 12 = 240$ chambers, so the remaining 10 DTs are found in station 4, sectors 4 and 10, in each wheel (Fig. 4-19).

Each tube cross section is $13 \times 42 \text{ mm}^2$, within which a single anode wire is made of gold-plated stainless-steel that has a diameter of $50 \mu\text{m}$. Different voltages are applied to the

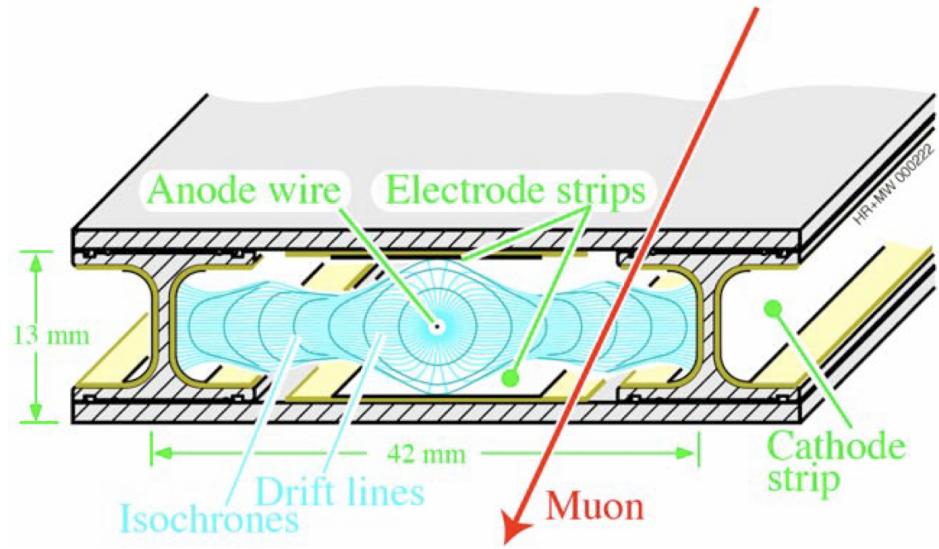


Figure 4-18. A cross section of a DT cell showing the drift lines (light blue), isochrone lines (dark blue), dimensions of the cell, and locations of the anode wire and cathode strips.

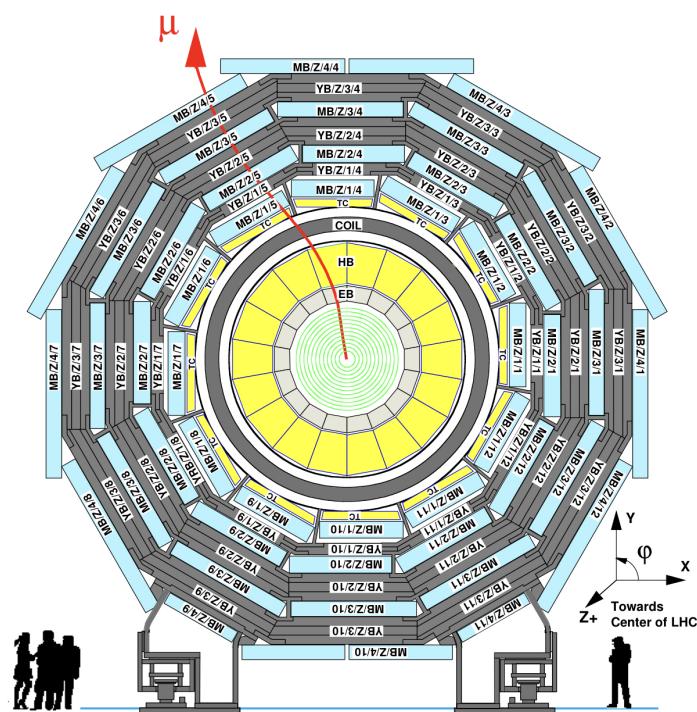


Figure 4-19. A cross section of CMS showing the locations and numbering scheme of the drift tubes in the barrel.

anode wires (+3600 V), electrode strips (+1800 V), and cathode strips (−1200 V). The gas mixture within a DT is similar to that of a CSC, containing approximately 85% Ar + 15% CO₂.

Physics: DTs are a good choice for the barrel because of the low rate and low strength of magnetic field. Uniform drift field of 1.5 KV/cm. A high p_T muon track will cross all 4 DT stations within the pseudorapidity range of $|\eta| < 0.8$. The reconstruction efficiency of such a track is better than 95%.

A single SL has a time resolution of only a few nanoseconds which provides excellent bunch crossing identification. To assist in determining muon p_T and timing, the DT system delivers the following muon track segment information to the L1 trigger: the position of the center of gravity (to a precision of 1.5 mm) and the corresponding angle (to a precision of 20 mrad), w.r.t. the SL reference frame. The total resolution of a DT in $r\text{-}\phi$ is about 100 μm which is comparable to the deviation caused by multiple scattering, for a muon with $p_T \leq 200$ GeV.

The gas gain is comparable to that of a CSC (about 100,000).

The maximum drift distance in a DT is 21 mm, as can be seen in Fig. 4-18, which corresponds to a drift time of 380 ns. This value

4.4.3 Resistive Plate Chambers

Overview: Even though CSCs 4.4.1 and DTs 4.4.2 cover the entire pseudorapidity range of $0 < |\eta| < 2.4$, redundancy is important. Therefore, resistive plate chambers (RPCs) are placed in both the endcaps and barrel to provide excellent timing and spatial resolution, comparable to that of scintillators, to supplement the measurements of the CSC and DT systems. An RPC is a gaseous parallel-plate detector

RPC has better time resolution than the 25 ns between two contiguous pp bunch crossings.

Design:

Gas mixture comprised of 96.2% C₂H₂F₄ (1,1,1,2-tetrafluoroethane) + 3.5% iC₄H₁₀ (isobutane) +

The RPC system consists of 480 rectangular chambers that are oriented parallel to the beam pipe. A single RPC is 2.455 mm long

Physics:

4.4.4 Gas Electron Multipliers

Overview:

Design:

Physics:

(GEM)

Located in the forward region. Filled with Ar/CO₂ gas mixture. The first GEMs At the time of this dissertation writing,

CHAPTER 5 THE COMPACT MUON SOLENOID EXPERIMENT

The proton-proton collisions at the LHC produce 100M pixels, 40M pictures per second. Only record 1K events/s some events only occur only once per 10B collisions!

Higgs bosons produced: 2010: 776 2011: 105K 2012: 500K? 2016: 2M Higgs, 30 Higgs/min 2017: 2.5M 2018: 3.5M

Use the GIF of the m4l plot with growing Data.

5.1 Trigger System

5.1.1 The Level-1 Trigger

Trigger system helps to manage the enormous collision rate at Size of 1 event is about 3 MB. This would be TB of data per second, which isn't feasible. Two stages of a trigger system: Level 1 Trigger (L1) and the High Level Trigger (HLT).

L1 Trigger has about a trigger rate of 100 KHz as a hard limit. It's latency is only 3.2 microsec.

At the start of a fill in the LHC, the trigger rate is maximum. INSERT L1 trigger rate plot.

In order to remain at the maximal L1 trigger, prescales are implemented which switch on new triggers, while turning off others. Collision data which passes certain triggers can be “parked” on disk and analyzed later (FIXME e.g., during Long Shutdowns) Data parking is used to Data scouting: The HLT has a trigger rate of about 1.5 KHz. It can write out about 5 GB/s to disk and has a much higher latency of 0.5 s compared to the L1 Trigger.

5.1.2 The High-Level Trigger

5.2 Event Reconstruction

5.2.1 Track and Vertex Reconstruction

5.2.2 Electron and Photon Reconstruction

5.2.3 Jet Reconstruction

5.2.3.1 Jet Energy Correction

5.2.3.2 Tagging b-Jets

5.2.4 Muon Reconstruction

5.2.5 MET Reconstruction

5.2.6 Tau Reconstruction

5.3 Event Simulation

5.3.1 Hard Interactions

5.3.2 Parton Showering and Hadronization

5.3.3 Detector Simulation

CHAPTER 6

HIGGS BOSON MASS MEASUREMENT IN THE $H \rightarrow ZZ^* \rightarrow 4\ell$ CHANNEL

6.1 Introduction

The Higgs boson was discovered in 2012 by the CMS and ATLAS collaborations. This was a momentous achievement in particle physics because the existence of the Higgs boson was required to complete the SM. In fact, it is sometimes referred to as the “missing puzzle piece” of the SM. The Higgs boson is one of a kind: it is the only fundamental scalar particle ever discovered so far. The unique boson could be a portal to new physics (*beyond Standard Model physics*, BSM), e.g., by decaying into BSM low-mass dilepton mass resonances (Chapter 7). In order to be certain that the recently discovered Higgs boson is truly the same as the one predicted by the SM, it is necessary to compare its measured properties to the predicted ones.

Some properties of the Higgs boson can be predicted by the SM, like - There are many results on Higgs properties: spin, charge, decay processes, lifetime, mass. - The last of these is the focus of this dissertation and is of particular importance to the Universe: depending on m_H and m_{top} , the stability of the Universe.

ALL previous mass measurements: - Run 1: - $H \rightarrow 2\gamma$ VALUE - $H \rightarrow ZZ \rightarrow 4L$ VALUE - Run 2: - $H \rightarrow 2\gamma$ VALUE - (2016) $H \rightarrow ZZ \rightarrow 4L$ VALUE - $H \rightarrow bb$ - $H \rightarrow \mu\mu$ - $H \rightarrow WW$

- Why this thesis is important: - This thesis describes the methodology and results of the best precision measurement of m_H to date by using the $hZZ4l$ decay and Full Run 2 data set from CMS. - Run 2 provides more data -> more precision on measurements of Higgs properties. - In addition to more $HZZ4l$ events, this analysis provides new techniques, specifically the VX constraint. - Predict m_H for Run 3, will start soon summer 2022 and provide an approximate 300? /fb of L int. - In 2026(?), HLLHC provides even more data. ref snowmass paper.

This chapter is structured as follows:

- General overview of the Higgs boson mass measurement ingredients (Section 6.2).
- Data sets, simulation, triggers (Section ??).
- Event reconstruction and selection (Section 6.4).
- Background estimation (Irred. and Reduc. Backgrounds).
- Signal modeling: kinematic discriminant, per-event mass uncertainties, VXBS constraint, reference to ad hoc studies in appendix.

- Systematic uncertainties.
- Results.
- Summary.

SEEMS TO BE A GOOD INTRO. Should it be the intro for the entire thesis?

6.2 Analysis Overview

The first step to performing a precision measurement of the Higgs boson mass (m_H) is to “observe” many Higgs bosons. However, production of a Higgs boson is essentially nonexistent in everyday conditions and is still extremely rare even in the high-energy pp collisions of the LHC. At a center-of-mass energy of 13 TeV, the total inclusive inelastic cross section of two protons colliding is 70 mb TODO: CITE. Comparing this to the production cross section of a Higgs boson (TODO $\sigma(pp\rightarrow H) = 59 \text{ pb}$) shows that a Higgs boson is produced in approximately one out of every billion pp collisions. TODO CITE

To complicate matters further, the Higgs boson has a *very* short mean lifetime of only $1.6 \times 10^{-22} \text{ s}$ [2]. Thus, the Higgs boson is not directly detected by CMS but is instead *inferred* from its stable decay products that enter the various subdetectors. Among all the fundamental particles so far discovered, the Higgs boson bears the second heaviest mass (approximately 125 GeV), the first belonging to the top quark (Section 2.1). This gives the scalar boson sufficient energy to decay into at least 9 different final states. **MENTION THAT NOT ALL DECAYS MAKE ON-SHELL PARTICLES?** Each decay occurs with a different probability—the *branching fraction* or *branching ratio* (\mathcal{B})—whose value depends on m_H as shown in Figure 6-1. The question then becomes, “*Which decay mode of the Higgs boson is most useful for the measurement of m_H ?*”. Owing to its large signal-to-background ratio of approximately 2 and its relatively rare four-lepton final state, the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel is selected and is called the *signal process*. Thus, a single Higgs boson decays via the signal process into two Z bosons (one on-shell and one off-shell) only 2.6% of the time. In turn, each Z boson decays into two opposite-sign, same flavor (OSSF) leptons ($Z \rightarrow \ell^+\ell^-$, where $\ell = e, \mu$) on average approximately

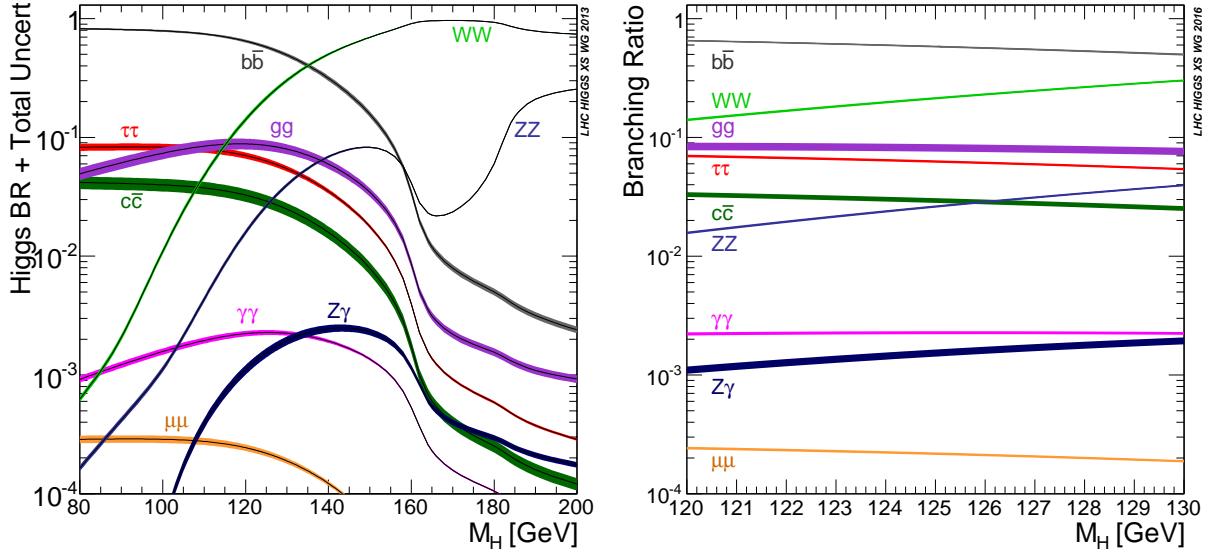


Figure 6-1. The branching ratios of various Higgs boson decays as a function of the Higgs boson mass over a wide range of values (Left) and a narrow range (Right).

6.7% of the time, giving rise to four distinct final states: $4e$, 4μ , $2e2\mu$, $2\mu2e$. The branching ratio for the overall signal process is then calculated as:

$$\mathcal{B}(H \rightarrow ZZ^* \rightarrow 4\ell) = \mathcal{B}(H \rightarrow ZZ^*) [\mathcal{B}(Z \rightarrow \ell^+\ell^-)]^2 = 1.8 \times 10^{-3}.$$

Thus, a signal event is expected to be produced only once in about every *trillion* pp collisions.

The strategy is then to search the pp collision data collected and analyzed by the CMS detector (Chapter 4) for all the detected $H \rightarrow ZZ^* \rightarrow 4\ell$ events. The task is not so straightforward; events in the data are categorized—not by the entire decay process—but by their final state, based on which triggers fired to collect which events. Section 6.3 describes the triggers used for this analysis to select events with the 4ℓ final state found in the corresponding data sets. The chosen events have a plethora of information from all the various subdetectors in CMS (Chapter 4) which gets reconstructed into *objects*—representations of the true particles produced by the event. Since the process of interest is $H \rightarrow ZZ^* \rightarrow 4\ell$, the reconstructed objects per event should be assembled in a fashion that coincides with the logic of the process. The reconstructed objects per event must then be assembled in a fashion that coincides with the logic of the process

of interest: $H \rightarrow ZZ^* \rightarrow 4\ell$. So, for example, two OSSF-dilepton objects should each appear to come from a Z-boson-like object (e.g. having a nominal mass of approximately 91 GeV and zero net electric charge)—instead of, say, coming from a Higgs-boson-like object. Furthermore, the involved objects should obey physics conservation laws (energy, momentum, charge, etc.) and pass detector selection criteria. These requirements are analysis-specific *event selection* process to ensure that the events

Once all signal-like events have been identified...

BACKGROUND

by being combined in a specific fashion and be combined in such a way so as to resemble this process, e.g.

BACKGROUND - By collecting events with the 4ℓ final state, we are likely to find signal events. - It's not just the signal process which produces 4ℓ : background also makes 4ℓ (Section FIXME).

- Before analyzing the data, however, it is important to make predictions using simulated samples (Section FIXME). - In order to sort signal from background, use simulated samples , which is the formation of particle physics objects from data. The data collected and analyzed by CMS is not so simple so as to have $H \rightarrow ZZ^*$

This process hinges on the conservation of momentum, since in the longitudinal (z) direction the pp collision has initial and final. Specifically, the - The Z boson has a precisely measured mass of TODO a neutral particle, so the two leptons into which it decays should combine to Group two leptons together, - Form two different pairs of opposite-sign, same-flavor (OSSF) leptons - If it appears that the to select specific hzz4l events (*event selection*).

6.3 Data Sets, Simulated Samples, and Triggers

6.3.1 Data Sets

6.3.2 Simulated Samples

6.3.3 Triggers

6.4 Event Reconstruction and Selection

6.4.1 Event Reconstruction

6.4.2 Event Selection

6.4.2.1 ZZ Candidate Selection

6.5 Background estimation

Measurement of the Higgs boson mass requires the accurate modelling of the total event yield in the signal region (SR), into which events can be categorized as either *signal* or *background*. These background events pass the signal event selection (Section 6.4.2.1) and thus spoil the purity of the signal events. This introduces further uncertainty into the final Higgs boson mass measurement. Therefore, it is a priority to both reduce and predict the expected number of background events, which can be split into *irreducible backgrounds* (IB) and *reducible background* (RB) processes.

6.5.1 Irreducible background

IB processes produce two Z bosons and each Z subsequently decays into two prompt leptons (leptons that emerge directly from the primary vertex). This reliably produces a 4ℓ final state, whose prompt leptons typically get reconstructed as four leptons that pass tight selection (*PTS leptons*), as defined in Sections ?? and ?? . The IB event then looks indistinguishable from the 4 PTS lepton of the *signal* process and cannot be reduced; Thus, throwing away IB process could mean throwing away a signal event. Since IB event cannot be reduced, it's called irreducible background. The two IBs for the HZZ4L analysis are:

- $gg \rightarrow ZZ \rightarrow 4\ell$ (gluon-gluon fusion),
- $q\bar{q}/gg \rightarrow ZZ \rightarrow 4\ell$ (quark-antiquark annihilation).

6.5.2 Reducible background

While irreducible background processes produce 4 prompt leptons, RB processes produce varying numbers of prompt and nonprompt leptons. Nonprompt leptons emerge from 3 main sources:

- misidentifying light-flavor hadrons (e.g., π^\pm) as leptons,
- heavy-flavor hadrons that decay mid-flight into leptons,
- and the asymmetric conversion of photons into electrons.

This analysis is tailored to efficiently reconstruct prompt leptons as those which pass tight selection (*PTS leptons*), whereas nonprompt leptons typically fail tight selection (*FTS leptons*). Tight and loose lepton selections are defined in Sections ?? and ???. Ideally, the SR would contain only the events with 4 prompt leptons, which would always be tagged as 4 PTS leptons. However, sometimes a truly nonprompt lepton is misidentified as a PTS lepton (sometimes called a *fake lepton*), depending on the kinematical properties of the lepton. This misidentification rate (f , sometimes called the *fake rate*) is due to imperfect detector performance, inefficiencies in reconstruction, and the specific choice of lepton selections used in the analysis.

If an event produces prompt and nonprompt leptons but is reconstructed as 4 PTS leptons (a 4P event), then it contaminates the SR. These processes constitute the RB processes (sometimes called “Z + X”). Examples of RB processes for this analysis include:

- Z + jets (yields 2 prompt leptons)
- $t\bar{t}$ + jets (yields 2 prompt leptons)
- WZ + jets (yields 3 prompt leptons)
- Z(Z/ γ^*) + jets (yields 4 prompt leptons).

The careful estimation of these RB contributions to the 4P region is necessary for the precise measurement of the Higgs boson mass.

6.5.2.1 OS Method

The goal of the OS Method is to estimate the number of opposite-sign same-flavor (OSSF) 4ℓ events produced by RB process that “contaminate” the 4P region (N_{4P}^{RB}), given by Eq. 6-13. However, the typical approach of using simulated samples does not model RB well, since RB processes (e.g., $Z + \text{jets}$) rely on higher-order effects like jet modelling which are not yet accurately simulated. Instead, a data-driven approach is used.

The logic of the OS Method is to study events in data that are similar to, but not exactly the same as, those found in the 4P region. Thus, events in data are sorted into 2 control regions (CRs), both of which are orthogonal to the 4P region and to each other:

- the $3P1F$ CR (built from 3 PTS and 1 FTS leptons)
- the $2P2F$ CR (built from 2 PTS and 2 FTS leptons).

The event selection for the 2P2F and 3P1F CRs is almost identical to that of the SR (Section 6.4.2.1), except that the FTS lepton(s) are required to build the Z_2 . The events that contribute mostly to the 2P2F(3P1F) CR are those that produce 2(3) prompt and 2(1) nonprompt leptons and are called $2pr(3pr)$ events. Similarly, the events that contribute mostly to the 4P SR are those that produce 4 prompt leptons and are called $4pr$ events.

The final formula for N_{4P}^{RB} is obtained by first supposing that event k is a $2pr$ event and contributes to the 2P2F CR an event weight of $w_{2pr \rightarrow 2P2F}^k$. This weight is built from the product of analysis weights (pileup, L1 pre-firing, etc., whose product is \hat{w}^k) and the reconstruction efficiencies (ϵ) of each lepton (ℓ_n):

$$w_{2pr \rightarrow 2P2F}^k = \hat{w}^k \cdot \epsilon_P^{pr}(\ell_1^k) \cdot \epsilon_P^{pr}(\ell_2^k) \cdot \epsilon_F^{np}(\ell_3^k) \cdot \epsilon_F^{np}(\ell_4^k), \quad (6-1)$$

where the superscript of ϵ indicates the lepton promptness (pr = prompt, np = nonprompt), the subscript indicates the lepton tightness status (P = PTS, F = FTS). To simplify the equations that follow, \hat{w}^k is set to unity. If the reconstruction efficiencies of a particular category are the same for

all j leptons across all k events (e.g., $\epsilon_P^{\text{pr}}(\ell_j^k) \equiv \epsilon_P^{\text{pr}}$), then Eq. 6-1 reduces to:

$$w_{2\text{pr} \rightarrow 2\text{P2F}}^k = \left(\epsilon_P^{\text{pr}} \right)^2 \left(\epsilon_F^{\text{np}} \right)^2. \quad (6-2)$$

Although a 2pr event mostly contributes to the 2P2F CR, a nonprompt lepton may be misidentified as a PTS lepton, depending on ϵ_P^{np} . Such an event would then fall into the 3P1F CR and, allowing for only one nonprompt PTS lepton at a time, contributes an effective weight of

$$\begin{aligned} w_{2\text{pr} \rightarrow 3\text{P1F}}^k &= \left(\epsilon_P^{\text{pr}} \right)^2 \left[\epsilon_P^{\text{np}}(\ell_1^k) \cdot \epsilon_F^{\text{np}}(\ell_2^k) + \epsilon_F^{\text{np}}(\ell_1^k) \cdot \epsilon_P^{\text{np}}(\ell_2^k) \right] \\ &= \left(\epsilon_P^{\text{pr}} \right)^2 [2\epsilon_P^{\text{np}}\epsilon_F^{\text{np}}]. \end{aligned} \quad (6-3)$$

Using the fact that a (non)prompt lepton is exclusively either PTS or FTS ($\epsilon_P^{\text{pr(np)}} + \epsilon_F^{\text{pr(np)}} = 1$), while recognizing that $\epsilon_P^{\text{np}} \equiv f$ (Section 6.5.2) and defining $\epsilon_P^{\text{pr}} \equiv \epsilon$, allows Eq. 6-3 to be written as

$$w_{2\text{pr} \rightarrow 3\text{P1F}}^k = 2\epsilon^2 f(1-f). \quad (6-4)$$

The prediction of Eq. 6-4 can be seen in Figures 6-6–6-8 for 2016–2018 UL data.

Even more rarely, both nonprompt leptons from a 2pr event may be misidentified as PTS leptons. In this case, event k contributes to the 4P region an effective weight of

$$w_{2\text{pr} \rightarrow 4\text{P}}^k = \epsilon^2 f^2, \quad (6-5)$$

where it is assumed that both leptons have the same misidentification rate. Similar equations can be derived for the contributions of a 3pr event to the 3P1F CR and to the 4P region:

$$w_{3\text{pr} \rightarrow 3\text{P1F}}^k = \epsilon^3 (1-f) \quad (6-6)$$

$$w_{3\text{pr} \rightarrow 4\text{P}}^k = \epsilon^3 f. \quad (6-7)$$

Since a 3pr event needs only 1 nonprompt PTS lepton to be included in the 4P region (therefore, carrying only 1 factor of f), a 3pr event tends to contribute more weight to 4P than does a 2pr event (which carries f^2).

If the total number of 2pr, 3pr, and 4pr events is $X_{2\text{pr}}$, $X_{3\text{pr}}$, and $X_{4\text{pr}}$, respectively, then

Figure 6-2 shows how the weight of a single event, derived in the previous equations (and others forthcoming), from each category contributes to the final yield of each CR ($N_{2\text{P}2\text{F}}, N_{3\text{P}1\text{F}}$) and of the SR ($N_{4\text{P}}$). It is then straightforward to evaluate the expected number of RB 4P events:

$$\begin{aligned} N_{4\text{P}}^{\text{RB}} &= \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 4\text{P}}^k + \sum_{m=1}^{X_{3\text{pr}}} w_{3\text{pr} \rightarrow 4\text{P}}^m \\ &= \sum_{k=1}^{X_{2\text{pr}}} \epsilon^2 f^2 + \sum_{m=1}^{X_{3\text{pr}}} \epsilon^3 f \\ &= f^2 \epsilon^2 X_{2\text{pr}} + f \epsilon^3 X_{3\text{pr}}, \end{aligned} \quad (6-8)$$

where only the quantities f and ϵ are known, so $X_{2\text{pr}}$ and $X_{3\text{pr}}$ must be estimated. This is achieved by relating $X_{2\text{pr}}$ to $N_{2\text{P}2\text{F}}$ using Eq. 6-2 across all 2pr events:

$$\begin{aligned} N_{2\text{P}2\text{F}} &= \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 2\text{P}2\text{F}}^k \\ &= \sum_{k=1}^{X_{2\text{pr}}} \epsilon^2 (1-f)^2 \\ &= (1-f)^2 \epsilon^2 X_{2\text{pr}} \end{aligned} \quad (6-9)$$

The strategy to relate $X_{3\text{pr}}$ to $N_{3\text{P}1\text{F}}$ is not as straightforward as relating $X_{2\text{pr}}$ to $N_{2\text{P}2\text{F}}$, since two other sources also contribute to the 3P1F CR (as shown in Figure 6-2):

- A 2pr RB process can yield one nonprompt PTS lepton (via Eq. 6-4).
- A 4-prompt IB process ($q\bar{q}/gg \rightarrow ZZ \rightarrow 4\ell$, “ZZ”) can yield one FTS lepton.

The second of these is well estimated from simulation, since ZZ produces 4 prompt leptons. If the total number of simulated ZZ events is $X_{4\text{pr}}^{\text{ZZ}}$, then event k belonging to these events contributes to the 3P1F CR an effective weight of

$$w_{4\text{pr}, \text{ZZ} \rightarrow 3\text{P}1\text{F}}^k = 4\epsilon^3 \epsilon_{\text{F}}^{\text{pr}}, \quad (6-10)$$

which accounts for any of the 4 prompt leptons to be reconstructed as a FTS lepton while the

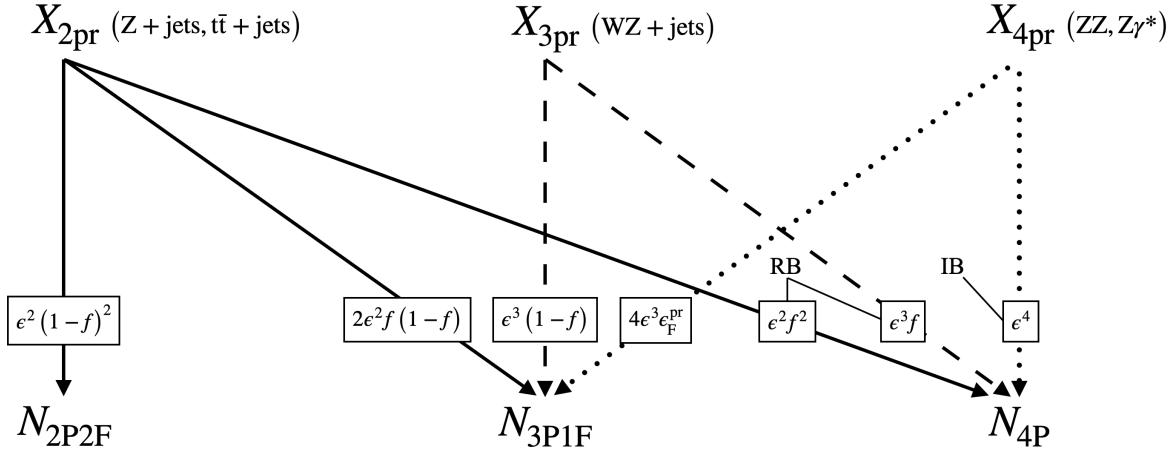


Figure 6-2. Contributions of per-event weights (boxed values) of various n -prompt-lepton processes ($X_{n\text{pr}}$, in parentheses) to the total numbers of events in the observed control regions ($N_{2\text{P}2\text{F}}$, $N_{3\text{P}1\text{F}}$) and signal region ($N_{4\text{P}}$). The labels “IB” and “RB” indicate those contributions which comprise the irreducible and reducible backgrounds, respectively.

others are PTS leptons. Incorporating Eqs. 6-4, 6-6, and 6-10 for all events relates $X_{3\text{pr}}$ to $N_{3\text{P}1\text{F}}$:

$$\begin{aligned}
 N_{3\text{P}1\text{F}} &= \sum_{j=1}^{X_{3\text{pr}}} w_{3\text{pr} \rightarrow 3\text{P}1\text{F}}^j + \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 3\text{P}1\text{F}}^k + \sum_{m=1}^{X_{4\text{pr}}^{\text{ZZ}}} w_{4\text{pr}, \text{ZZ} \rightarrow 3\text{P}1\text{F}}^m \\
 &= \sum_{j=1}^{X_{3\text{pr}}} \epsilon^3 (1-f) + \sum_{k=1}^{X_{2\text{pr}}} 2\epsilon^2 f (1-f) + \sum_{m=1}^{X_{4\text{pr}}^{\text{ZZ}}} 4\epsilon^3 \epsilon_F^{\text{pr}} \\
 &= (1-f)\epsilon^3 X_{3\text{pr}} + 2f(1-f)\epsilon^2 X_{2\text{pr}} + 4\epsilon^3 \epsilon_F^{\text{pr}} X_{4\text{pr}}^{\text{ZZ}} \\
 &= (1-f)\epsilon^3 X_{3\text{pr}} + 2f(1-f)\epsilon^2 X_{2\text{pr}} + N_{3\text{P}1\text{F}}^{\text{ZZ}}, \tag{6-11}
 \end{aligned}$$

where $N_{3\text{P}1\text{F}}^{\text{ZZ}}$ is simply the raw (integer) number of ZZ events that pass 3P1F selections, obtained directly from simulation.

At this point, $N_{4\text{P}}^{\text{RB}}$ can be isolated by combining Eqs. 6-8, 6-9, and 6-11:

$$N_{4\text{P}}^{\text{RB}} = \left(\frac{f}{1-f} \right) N_{3\text{P}1\text{F}}^{\text{Data}} - \left(\frac{f}{1-f} \right)^2 N_{2\text{P}2\text{F}}^{\text{Data}} - \left(\frac{f}{1-f} \right) N_{3\text{P}1\text{F}}^{\text{ZZ}}. \tag{6-12}$$

Thus, Eq. 6-12 estimates the RB contribution to the 4P region by using a single lepton misidentification rate to reweight the raw yields of events found in the 3P1F and 2P2F CRs of

data, and also the 3P1F CR of ZZ.

It should be mentioned that the above formula was derived assuming that the per event analysis weights were set to unity ($\hat{w}^k = 1$) before being scaled by the misidentification rates. It was also assumed that f is a constant, for all nonprompt leptons misidentified as PTS leptons across all events, which is not the case as is shown in Figure 6-9. Thus, an extension of Eq. 6-12 can be formed by assigning misidentification rates that depend on the kinematical variables per lepton, by restoring the analysis weights per event, and by summing over the total number of raw yields per CR:

$$N_{4P}^{RB} = \sum_{i=1}^{N_{3P1F}^{\text{Data}}} \hat{w}^i \frac{f_i}{1-f_i} - \sum_{j=1}^{N_{2P2F}^{\text{Data}}} \hat{w}^j \frac{f_{1,j}}{1-f_{1,j}} \frac{f_{2,j}}{1-f_{2,j}} - \sum_{k=1}^{N_{3P1F}^{\text{ZZ}}} \hat{w}_{ZZ}^k \frac{f_k}{1-f_k}, \quad (6-13)$$

where $f_{1,j}$ and $f_{2,j}$ are the misidentification rates of the first and second FTS leptons, respectively, found in the j^{th} 2P2F event and \hat{w}_{ZZ}^k accounts for the differential QCD and electroweak k -factors defined in Section 6.5.1.

6.5.2.2 Lepton misidentification rate measurement

As mentioned in Section 6.5.2, the lepton misidentification rate (f) is the probability that a nonprompt lepton will pass tight selections (PTS). The value f is a function of the flavor ($\ell = e, \mu$), p_T , and η of a lepton. The misidentification rate is calculated by simply counting the number of nonprompt PTS leptons (N_p^{np}) that enter a particular $\ell, p_T, |\eta|$ bin compared to the total number of loose probe leptons (N_L^{np}) in the same bin:

$$f(\ell, |\eta|, p_T) = \frac{N_p^{\text{np}}}{N_L^{\text{np}}}. \quad (6-14)$$

The p_T^e bin edges are [5–10–20–30–40–50–80] GeV and the p_T^μ bin edges are [5–7–10–20–30–40–50–80] GeV. The nonprompt leptons used to measure f are taken from events in data with a signature like that of $Z + \ell_L$, where Z is a Z boson and ℓ_L is a loose lepton. By construction, this region of events is orthogonal to the 2P2F, 3P1F, and 4P regions, and provides a clean source of ℓ_L . The loose lepton, whose selection is defined in Sections ?? and ??, is also called the *probe* lepton. The probe lepton is either a PTS or FTS lepton and is counted

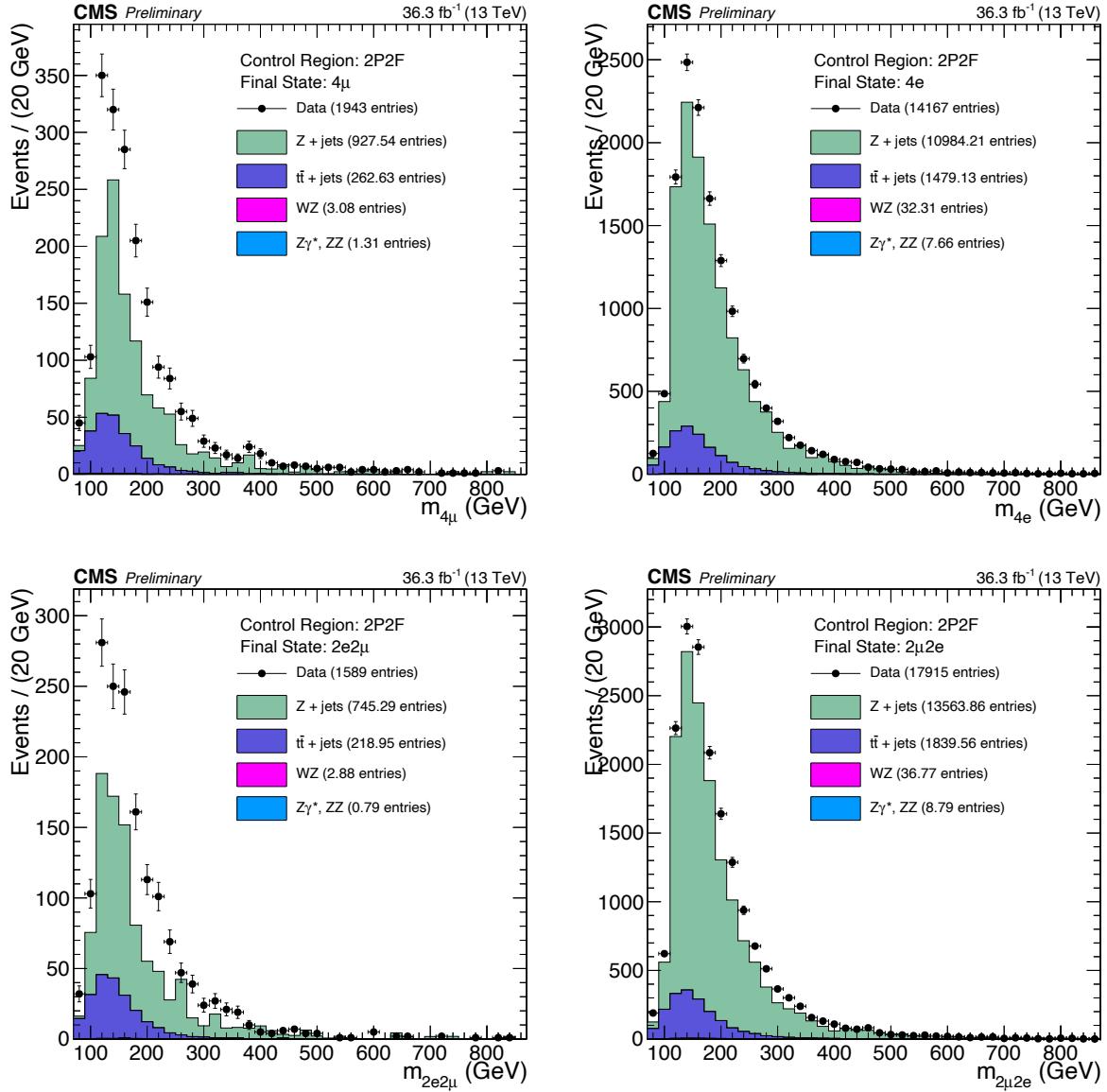


Figure 6-3. Events from 2016 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples (Z + jets, tt + jets, WZ, Z γ^* , ZZ). The results are split into the 4 ℓ final states: 4 μ (top left), 4e (top right), 2e2 μ (bottom left), 2 μ 2e (bottom right).

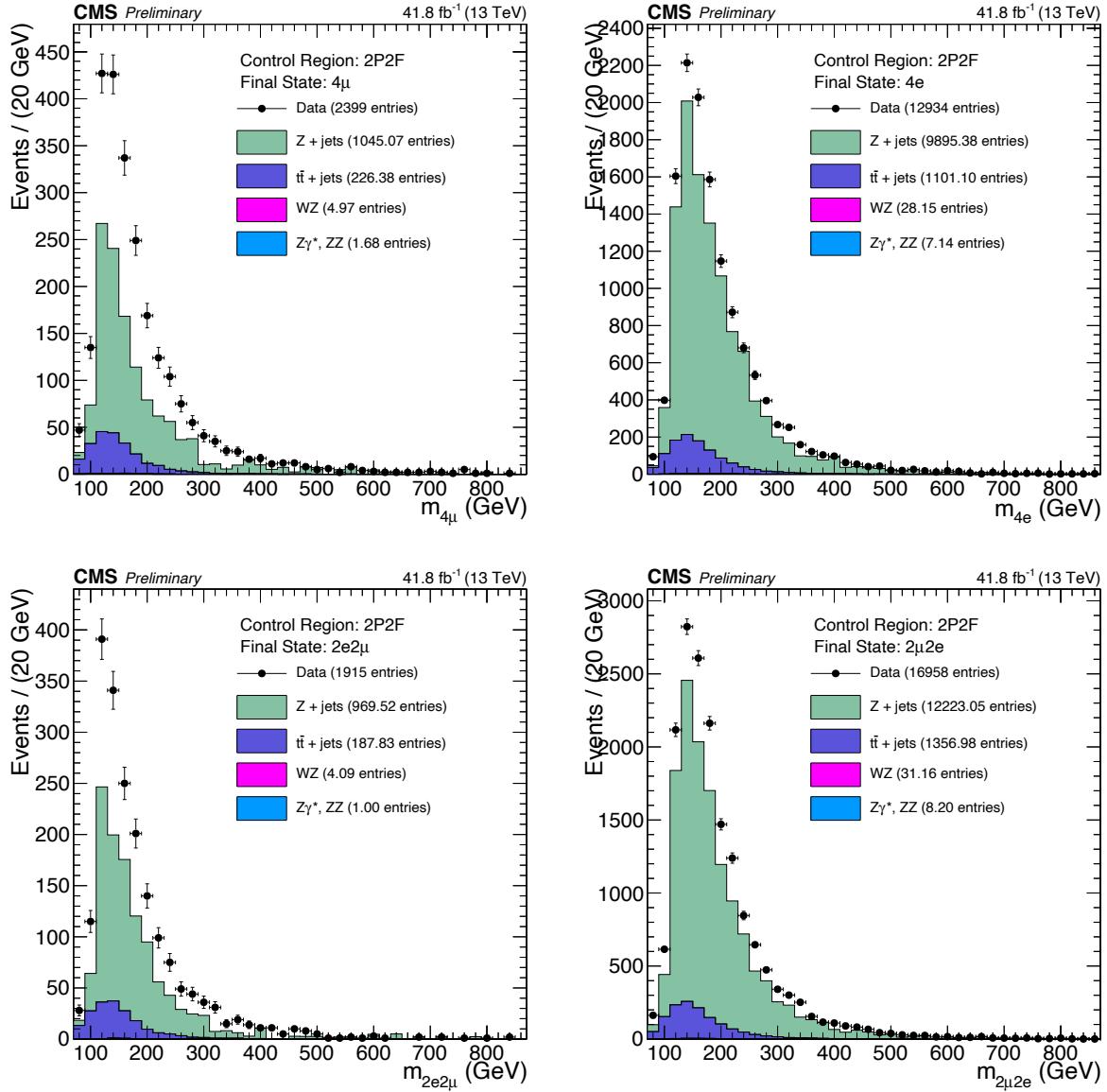


Figure 6-4. Events from 2017 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, Z/γ^* , ZZ). The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

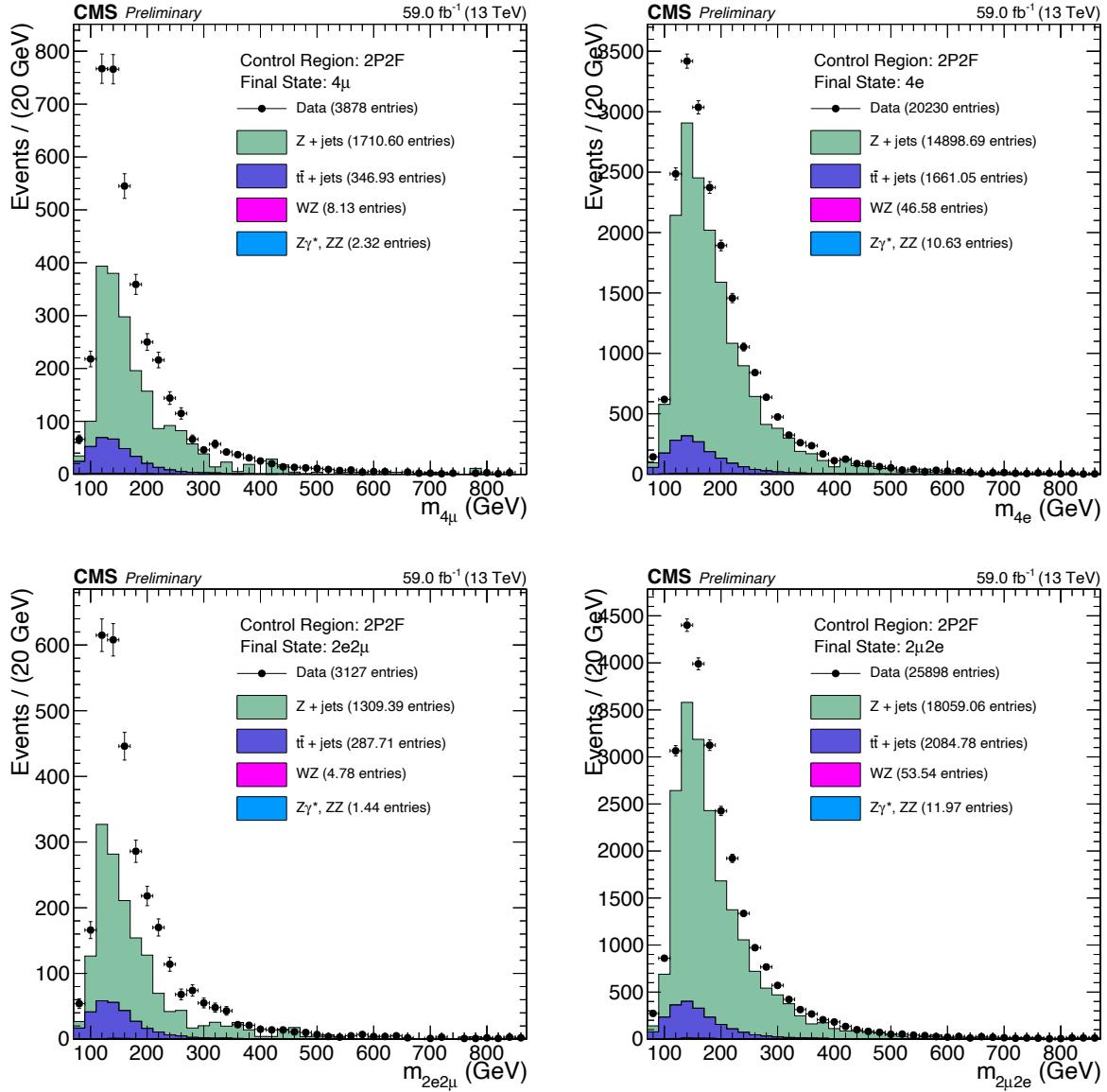


Figure 6-5. Events from 2018 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, Z/γ^* , ZZ). The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

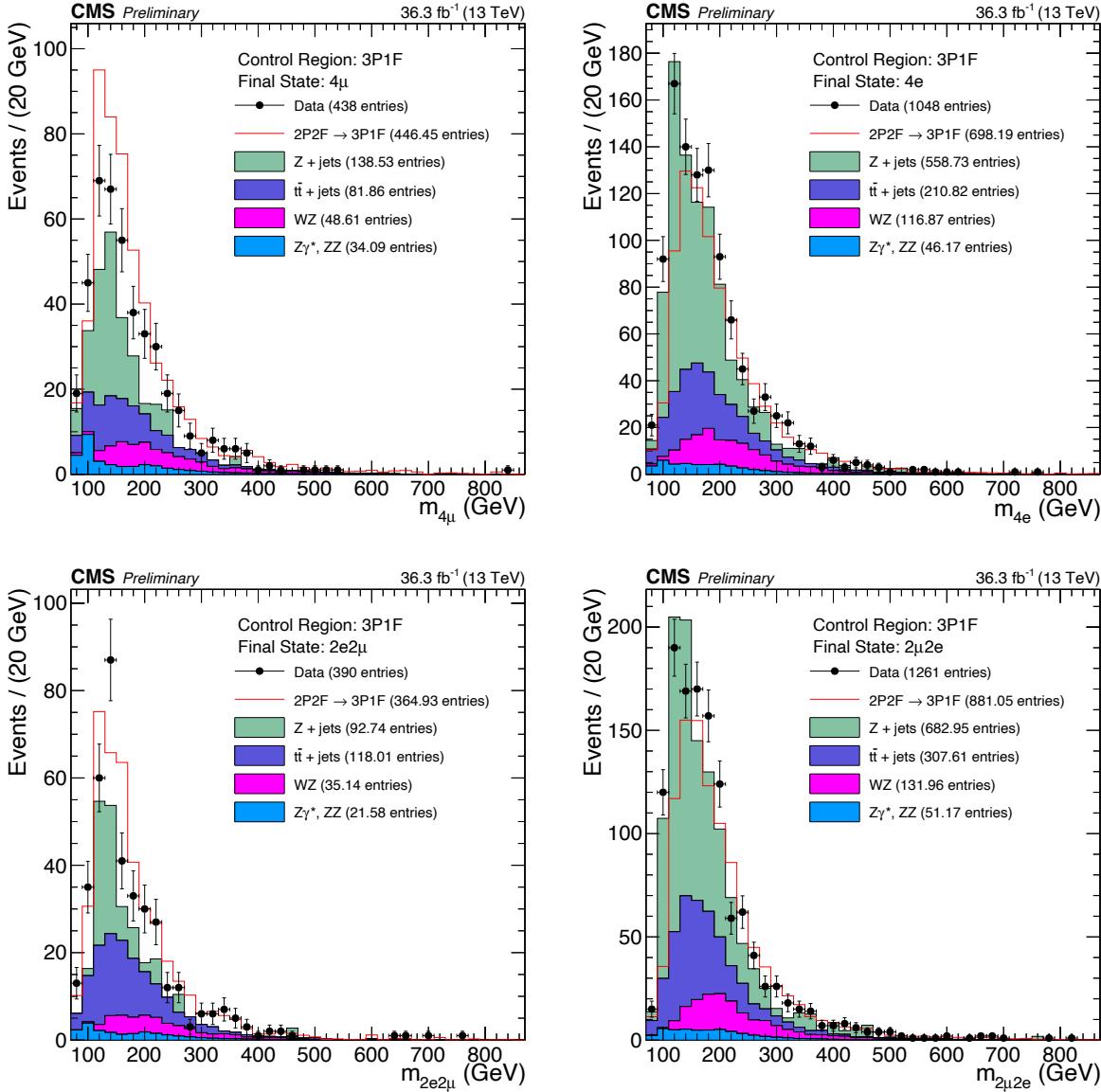


Figure 6-6. Events from 2016 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ , Z/γ^* , ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 6-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

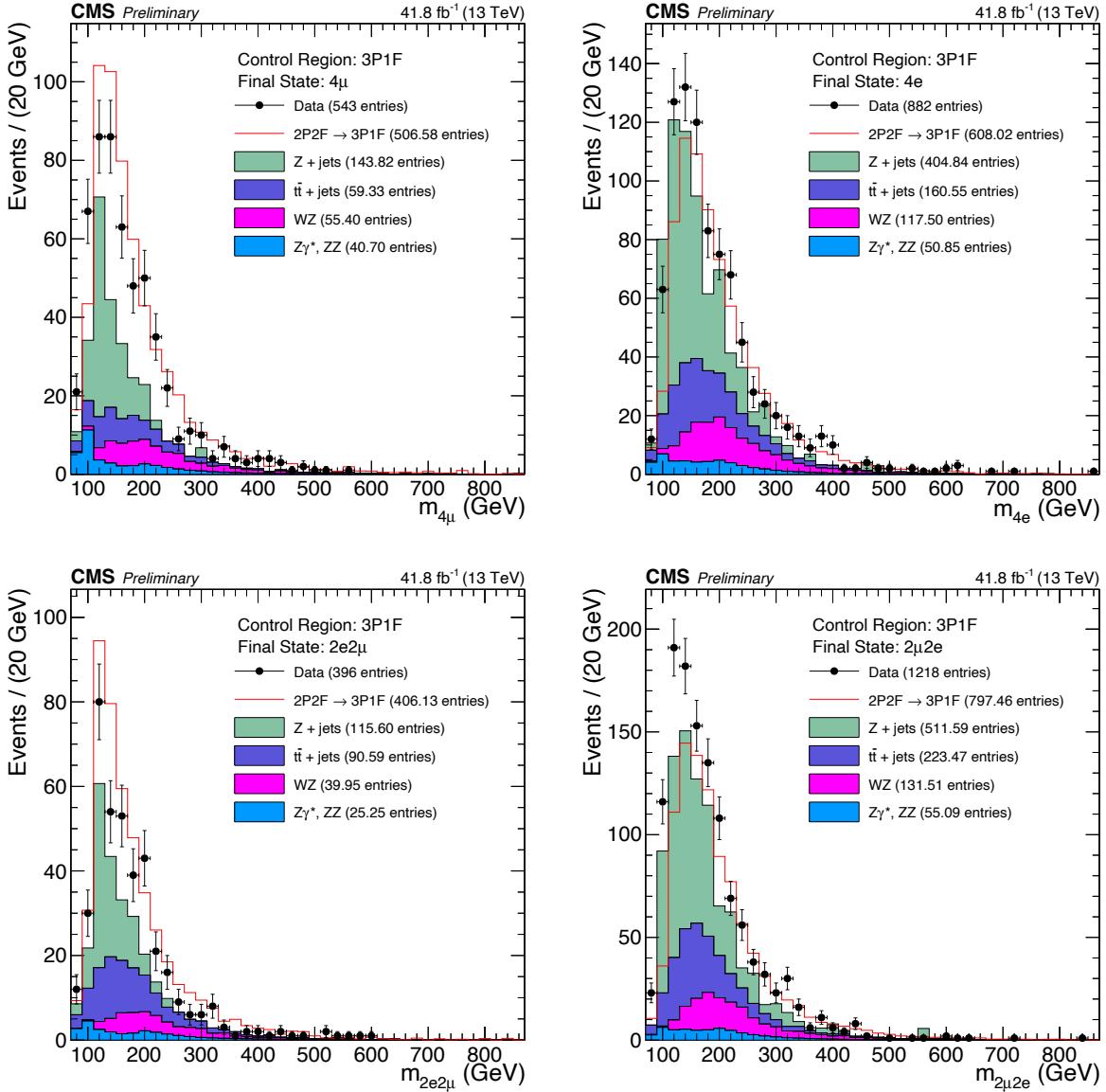


Figure 6-7. Events from 2017 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples (Z + jets, tt + jets, WZ, Z γ^* , ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 6-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4 ℓ final states: 4 μ (top left), 4e (top right), 2e2 μ (bottom left), 2 μ 2e (bottom right).

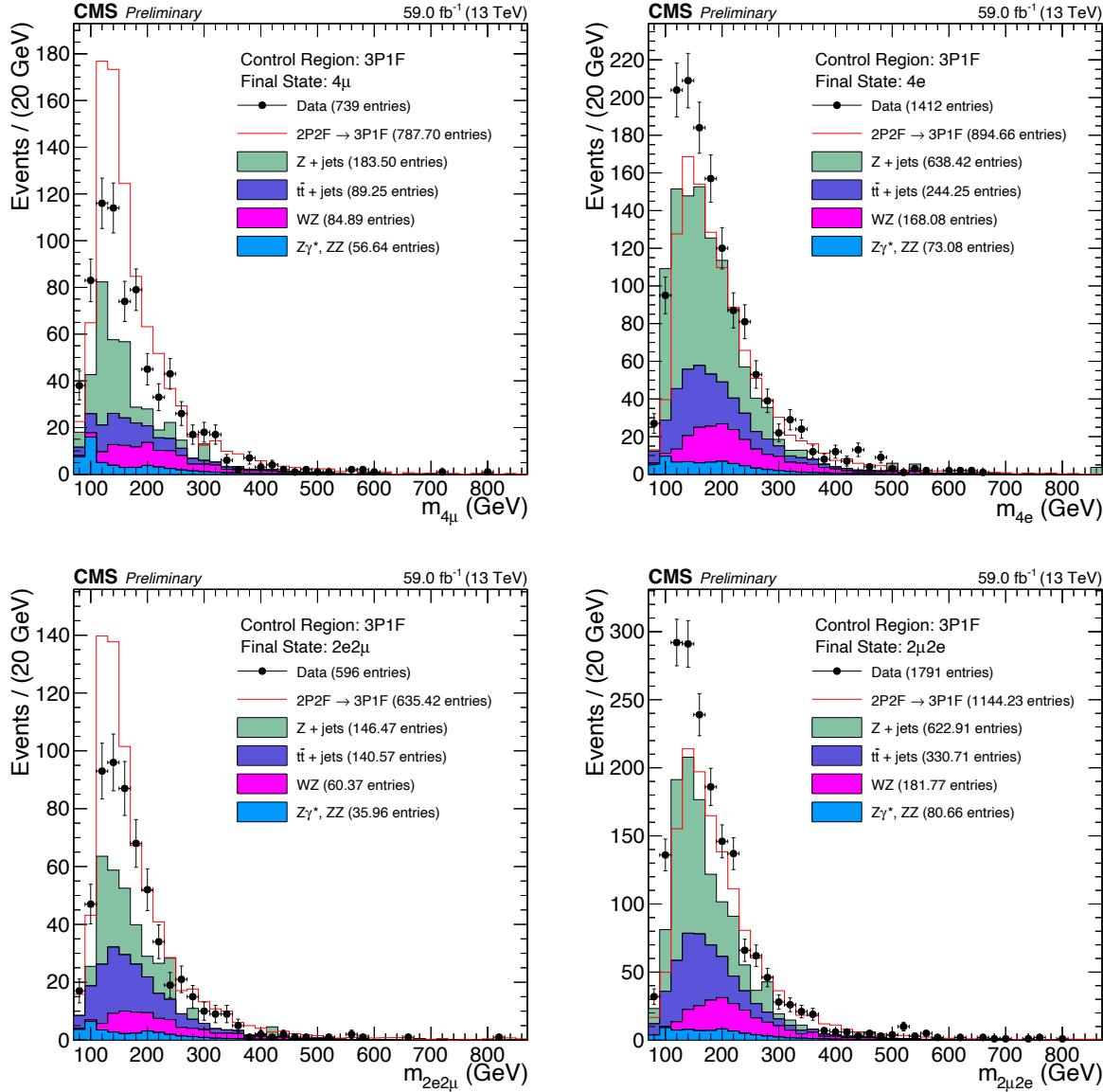


Figure 6-8. Events from 2018 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ , Z/γ^* , ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 6-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

towards both the numerator and denominator of Eq. 6-14.

Events are selected that satisfy the following criteria:

- The event has exactly 3 leptons.
- The event contains $E_T^{\text{miss}} < 25 \text{ GeV}$.
- Two of the leptons form a Z candidate. A Z candidate is formed when:
 - The lepton pair is OSSF.
 - Both leptons PTS.
 - The leading lepton has $p_T > 20 \text{ GeV}$.
 - The subleading lepton has $p_T > 10 \text{ GeV}$.
 - The lepton pair satisfies $|m_{\ell\ell} - m_{Z_{\text{PDG}}}| < 7 \text{ GeV}$
- The third and final lepton is loose (and may be a PTS or FTS lepton).
- Suppress QCD processes: probe lepton and OS lepton from Z have $m_{\ell\ell} > 4 \text{ GeV}$.

The calculation of f requires that ℓ_L is a nonprompt lepton but since data events were used, this may not be the case. For example, the decay of WZ produces 3 prompt leptons and so this contribution must be subtracted. Thus, the number of expected prompt probe leptons from WZ events is subtracted from both the numerator and denominator in Eq. 6-14 for each $\ell, p_T, |\eta|$ bin. The final OS Method misidentification rates for electrons and muons are shown in Figure 6-9 using 2016–2018 UL data.

6.6 Signal Modeling

6.6.1 Z_1 Mass Constraint

6.6.2 Per-Event Relative Mass Error Categorization

6.6.3 VXBS Constraint

6.6.4 Matrix Element-Based Kinematic Discriminant

Expected uncertainty	4μ	$4e$	$2e2\mu$	$2\mu2e$	Inclusive
1D (No bkg)	153	466	315	300	121

Table 6-1

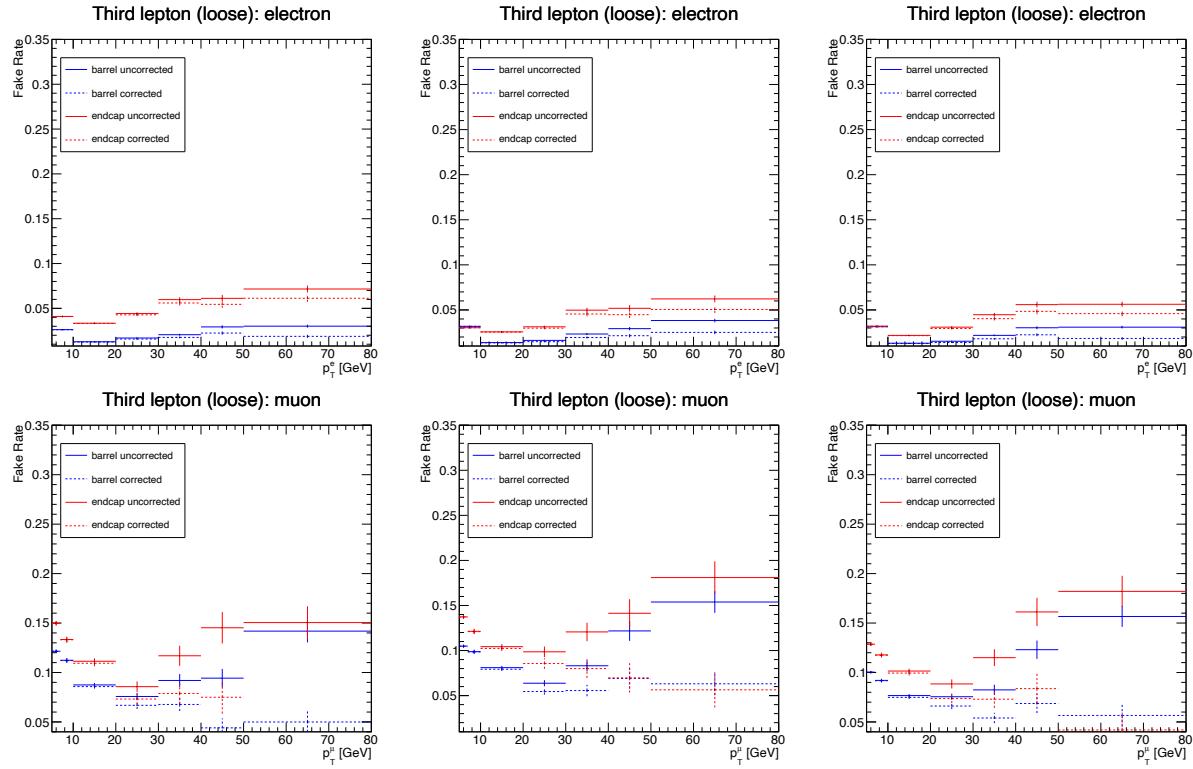


Figure 6-9. Electron (top row) and muon (bottom row) fake rates evaluated using the OS Method vs. the p_T of the probe lepton for 2016, 2017, and 2018 UL data (left, middle, right columns, respectively).

6.7 Systematic Uncertainties

Words.

6.8 Results

Words.

6.9 Summary

Words.

CHAPTER 7
SEARCH FOR LOW-MASS DILEPTON RESONANCES IN THE $H \rightarrow ZX/XX \rightarrow 4\ell$
CHANNEL

7.1 Introduction

Words.

CHAPTER 8 SUMMARY

words.

REFERENCES

- [1] CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(8):S08004, August 2008. doi: 10.1088/1748-0221/3/08/S08004. URL <https://ui.adsabs.harvard.edu/abs/2008JInst...3S8004C>.
- [2] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020. doi: 10.1093/ptep/ptaa104. and 2021 update.